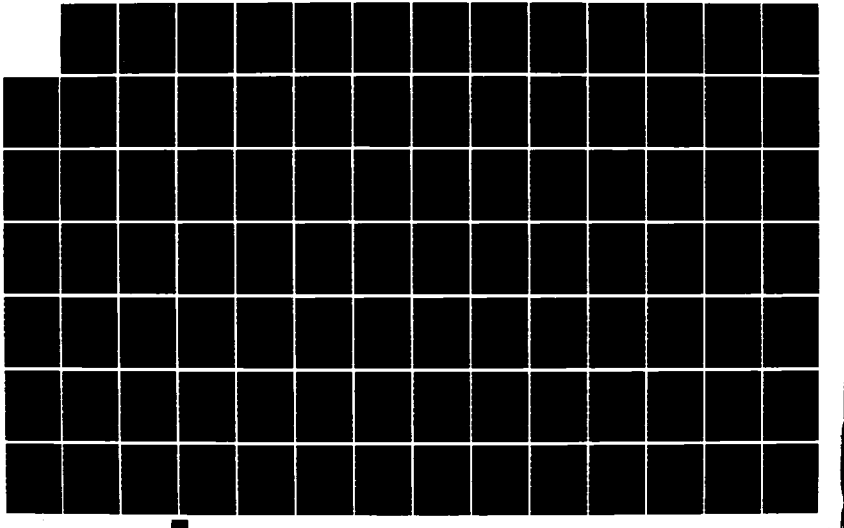
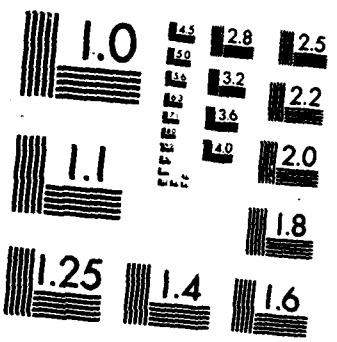
AD-A173 032 APPLICATION OF THE AM1 AND MNDO SEMIEMPIRICAL QUANTUM MECHANICAL MOLECULAR MODELS(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH K M DIETER DEC 86 1/2

UNCLASSIFIED AFIT/CI/NR-86-183D F/G 7/4 NL





SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFIT/CI/NR 86-183D	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* Application of the AM1 and MNDO Semiempirical Quantum Mechanical Molecular Models		5. TYPE OF REPORT & PERIOD COVERED THESIS/DISSERTATION
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Kenneth Michael Dieter		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS AFIT STUDENT AT: The University of Texas		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE 1986
		13. NUMBER OF PAGES 164
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* UNCLASS
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

APPROVED FOR PUBLIC RELEASE: IAW AFR 190-1

Lynn Wolaver 25 Jul 86

LYNN E. WOLAVER
Dean for Research and
Professional Development
AFIT/NR

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

ATTACHED ...

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

86 10 10 095

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

APPLICATION OF THE AM1 AND MNDO SEMIEMPIRICAL QUANTUM MECHANICAL MOLECULAR MODELS

BY

KENNETH MICHAEL DIETER, B.S., M.S.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

December, 1986

TO

SHARON

APPLICATION OF THE AM1 AND MNDO SEMIEMPIRICAL

QUANTUM MECHANICAL MOLECULAR MODELS

APPROVED BY
SUPERVISORY COMMITTEE:

Michael J. S. Dewar

S. E. Wells

F. A. Matsen

VITA

Kenneth Michael Dieter was born in Aurora, Illinois, on March 31, 1952, the son of Betty Jane Does and Kenneth Mathias Dieter. He was graduated from East High School of Aurora, Aurora, Illinois, in June 1970, and entered the United States Air Force Academy, Colorado. He received the degree of Bachelor of Science in Chemistry from the United States Air Force Academy in June, 1974 and was commissioned a second lieutenant in the United States Air Force. He entered Michigan State University that same year under the sponsorship of the United States Air Force, and received the degree of Master of Science from Michigan State University in March, 1975. He married the former Sharon Elaine VanValin in June, 1975. During the following years, he served as a missile maintenance officer at Whiteman AFB, Missouri, (1975-1979) and as an instructor of chemistry at the United States Air Force Academy, Colorado (1979-1983). During this period, his two daughters were born, Rebekah Anne (1977) and Sarah Marie (1979). In August, 1983, he entered the Graduate School of The University of Texas at Austin under the sponsorship of the United States Air Force.

Permanent Address: 735 Front Street
Aurora, Illinois 60505

ACKNOWLEDGEMENTS

My deep appreciation goes to Professor Michael J. S. Dewar, without whose support and encouragement, this challenge would have been insurmountable, and without whose infectious enthusiasm and creativity in his unending pursuit of understanding, it would have been tedious and unmotivating.

It would be impossible to fully express the part my wife, Sharon, and my daughters, Rebekah and Sarah, played in my reaching this point. We have sacrificed as one, and we have achieved as one.

I also want to thank my parents and Sharon's parents, as well as other friends and relatives, who have sustained us through their prayers.

To the members of the Dewar Group, I am indebted for their enlightening discussions, friendship and support, and particularly for their willingness to listen when I could not contain my frustrations any further.

I am also indebted to the United States Air Force and the Air Force Institute of Technology for sponsoring my graduate studies.

Finally, I want to acknowledge my total dependence on my Lord and Savior, Jesus Christ, through Whom all things are possible, and without Whom nothing is possible.

ABSTRACT

Studies were conducted involving the three phases of application of the AM1 and MNDO semiempirical quantum mechanical molecular models to problems in chemistry: development, testing, and use in specific investigations. As developmental work, efforts were made to obtain parameters for phosphorus valid for both trivalent and pentavalent compounds. Both the standard AM1 and MNDO algorithms, and algorithms including a core - core repulsion function modified to take into account the change in bonding interactions with change in valency, were used in these efforts. Contributory to the testing of the AM1 model, results for proton affinities and deprotonation enthalpies were extensively studied. The validity, as well as the limitations, of using AM1 in studies of reactions involving these processes were evaluated. Finally, MNDO and AM1 were used to examine two specific problems of chemical and theoretical interest, the condensation reaction of polyketide biosynthesis and the mass spectral fragmentation of cis-1-nitropropene. Results in the former study support the suggestion that a crucial factor in enzyme reactions is the exclusion of water from the reacting system. In the latter study, AM1 results were shown to correlate well with the fragmentation of the cis-1-nitropropene molecular ion.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

While man has always had a natural curiosity about the chemical processes surrounding him, the understanding of them grew slowly prior to the late 18th and early 19th centuries. Dalton's presentation of his theory of atoms in 1802, along with Lavoisier's discovery of mass as the fundamental quantity in chemical reactions in 1789, provided the basic concepts necessary to understand chemical reactions. Progress in the determination of chemical composition and the characteristic reactivity of elements and functional groups, and development of bonding theories and theories of reaction mechanisms followed. An explosion of instrumentation in the 20th century added to the momentum, providing information unobtainable from wet chemical techniques. Spectroscopic, crystallographic and diffraction techniques gave new insights into molecular structure, bond strength, electron distribution, group reactivity and stereochemistry. As a result, a greater depth of understanding of reactions was reached.

Still, there remained a realm of information unreachable by either wet chemical or instrumental techniques. While enthalpic and entropic properties of transition states could be derived from experiments, the structures of the transition states, and, hence, the elementary processes involved in reaching the critical points of the reactions, could not. Furthermore, this was not a result of insufficiently developed techniques, but of a fundamental limitation due to the uncertainty principle.[1]

Advances in theoretical calculations have opened new doors in this area. Calculations do not have the same limitations as experimental techniques. The properties of species corresponding to any point on a potential energy (PE) surface can be estimated, not just those of stable species. Thus, calculations can estimate the properties of transition states. The word "estimate" is required in this context, however, because theoretical calculations do have other limitations.

If the Schrödinger equation could be solved exactly for any system of interest, chemists would, in principle, be able to unequivocally explain any chemical process.[2] For anything other than the hydrogen atom or hydrogen molecular cation, this is not possible. As a result, calculations are simply one more tool to be used in concert with others to glean information concerning exactly what transpires during a reaction. Furthermore, to obtain this tool, approximations have to be made. Different quantum mechanical

models result from acceptance of varying sets of approximations.[1]. The acceptability of a model, i.e. a specific set of approximations, for a given study must be judged on the basis of two criteria: (1) the feasibility of applying the model in the study, and (2) the accuracy of the model in predicting properties for the types of species involved in the study.

The feasibility of using a model is most commonly determined by the time required for the calculations, a model being without value if the calculations take an unreasonable length of time. While some models, e.g. the Hückel model, can be applied to some problems without the use of computers, most models require computers for even the simplest studies. Even with computers, the use of some models is unfeasible or feasible only if further simplifications, such as symmetry or fixed geometry, are imposed. There must be, however, substantial evidence from experimental studies for these simplifications, or their effects must be thoroughly tested for and found to be negligible or predictable with respect to the aspects of the chemical process being investigated.[3] Otherwise, they will be self-defeating, enabling use of the model but casting doubt on the conclusions drawn from the calculations.

Even if its use in a particular study is feasible, a model is still of little value if it gives inaccurate results for the type of chemical process of interest. Since all models are based on sets of approximations, the accuracy of results cannot be assumed, but

must be established by comparison of calculated values of observables with experimental values.[1] The more properties of stable species estimated well by a specific model, the more likely it is that the model will provide good descriptions of species corresponding to other points on the PE surface.

The application of a quantum mechanical molecular model, then, involves three phases: (1) development of the set of approximations which define the model, (2) testing of the model, and (3) use of the model in the investigations of specific chemical problems. The following studies cover all three of these aspects with respect to the MNDO[4] and AM1[5] semiempirical models. The first study (Chapter 2) involves attempts to parameterize phosphorus for both normal and hypervalent compounds. The performance of AM1 in calculating proton affinities and deprotonation enthalpies is evaluated in the second study (Chapter 3) to determine the validity and limitations of using this model in studies of reactions involving protonation and deprotonation. In the last two studies, MNDO and AM1 are used to study the condensation reaction in polyketide biosynthesis (Chapter 4) and the mass spectral fragmentation of cis-1-nitropropene (Chapter 5).

References

1. Dewar, M. J. S. *J. Phys. Chem.* 1985, *89*, 2145.

2. Dewar, M. J. S. *J. Mol. Struct.* 1983, *100*, 41.

3. This will, of course, be particularly difficult for transition states for the reasons given above.

4. Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* 1977, *99*, 4989, 4907.

5. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E.F., Stewart, J. J. P. *J. Am. Chem. Soc.* 1985, *107*, 3902.

Chapter 2

PARAMETERIZATION OF MNDO AND AM1 FOR NORMAL AND HYPERVALENT PHOSPHORUS

Introduction

MINDO/3,[1] MNDO,[2] and AM1[3] are semiempirical quantum mechanical molecular models based on the intermediate neglect of differential overlap (INDO)[4,5] and neglect of diatomic differential overlap (NDDO)[4,6] approximations to the Roothaan-Hall equations,[7] which themselves are an approximation of the Hartree-Fock (HF) self-consistent field (SCF) equations.[8] While the terms INDO and NDDO refer basically to the neglect of various integrals in the solution of the Roothaan-Hall equations, the term "semiempirical" refers to the replacement of other integrals by parametric functions. The parameter values are selected to make the calculated values of a number of molecular properties for a wide range of molecules agree as closely as possible with experimental values.

The use of parametric functions with empirically derived parameters serves a three-fold purpose. First, by reproducing "reality", as represented by experimental data for a large number of molecules of varying types, an allowance for electron correlation, which exists in reality but is ignored in the HF/SCF method,[9] is

implicitly included in the calculations. Secondly, implicit compensation is also made for the neglected integrals . Thirdly, computational time is greatly reduced. All of these results play essential roles in attaining the goal behind the development of this series of models,[1,2,3] that is to provide a method of calculating molecular properties with sufficient accuracy and reliability, at low enough cost, to be of practical value in investigating "real life" chemical problems, particularly with regard to areas where experimental data are lacking or impossible to obtain (e.g. transition state geometries).[10]

Given that these models are based on numerous approximations, many of which are approximations to approximations, it is indeed surprising how successful they are. Among the properties for which MINDO/3, MNDO and AM1 have been shown to provide good estimates are heats of formation, [1,2,3] molecular geometries,[1,2,3] dipole moments,[1,2,3] ionization potentials,[2,3] electron affinities,[11] and proton affinities and deprotonation enthalpies.[12] For heats of formation, results are comparable with *ab initio* calculations with split valence basis sets.[10,13]

In general, the success of these methods extends beyond calculations involving the organic elements (carbon, hydrogen, oxygen and nitrogen). Other elements for which one or more methods have been parameterized include beryllium,[14] boron,[15] fluorine,[16] aluminum,[17] silicon,[18] phosphorus,[19] sulfur,[20] chlorine,[21]

bromine,[22] tin,[23] iodine,[24] mercury,[25] and lead.[26]

A limitation has been evident, however, in the ability of these methods to handle the so-called hypervalent compounds, i.e. compounds in which at least one atom has an expanded octet. For example, the average absolute error in the MNDO heats of formation for 10 trivalent phosphorus compounds (P_2, PN, P_4, PH_3, PF_3, PCl_3, $P(CH_3)_3$, $P(OCH_3)_3$, P_2H_4, P_4O_6) is 29.0 kcal/mol,[27] while that for 5 pentavalent compounds (PF_5, PCl_5, F_3PO, Cl_3PO, $(CH_3)_3PO$) is 76.9 kcal/mol.[27] Unlike the individual errors for the trivalent compounds, some of which are negative while others are positive, all errors for the pentavalent compounds are positive. The obvious conclusion from these results is that the bonding parameters do not adequately compensate for the increased electron-electron and core-core repulsion resulting from bonding five atoms instead of three to the phosphorus. Attempts to correct this deficiency by adding d-atomic orbitals (d-AO) in the CNDO approximation[29] to the basis set failed for sulfur[30] and phosphorus.[31] When the β_d value for sulfur was allowed to vary to compensate for contraction of the d-AOs with increasing charge on the hypervalent atom, results were much improved. Because of the functional dependence of the β_d parameter on atomic charge, however, the electronic wavefunctions were nonvariationally optimized. As a consequence, full SCF calculations were required for gradient calculations (see below). The additional computation time required due to this factor, plus that due to expansion of the basis set to include d-AOs, made the

calculations too slow to be of general use.[30]

Ab initio calculations have indicated that although d-AOs do not participate in the bonding of hypervalent phosphorus to the extent indicated in sp^3d^n hybridization models, they do increase the calculated stability of hypervalent molecules, and are required to accurately describe the geometries of the phosphine oxides.[32,33,34] Specifically, without d-AOs the P-O bond length of the hypothetical H_3PO molecule was calculated to be approximately the same as that of a P-O single bond, i.e. 1.6Å.[34] With the inclusion of d-AOs, the bond length decreased to 1.47Å,[34] which is close to the 1.48Å of $(CH_3)_3$ PO.[32] The inclusion of d-AOs, and even higher angular momentum functions, were also found to be essential to correctly predict the order of stability of P_2, P_4, and P_8.[34,35,36,37]

MINDO/3, MNDO and AM1, however, have the flexibility of parametric functions derived from experimental data, which the *ab initio* calculations, of course, do not. As a consequence, even without d-AOs, MNDO calculates P_4 as more stable than $2P_2$, and the P-O bonds in the phosphine oxides at approximately the correct length.[28] MNDO does calculate cubic P_8 to be 68 kcal/mol more stable than $2P_4$, however.[38]

Despite this last discrepancy, then, it appears the parametric functions in MNDO are sufficient to compensate for at least some of the factors for which d-AOs are required in *ab initio* calculations. Since addition of d-AOs in the CNDO approximation did

not give satisfactory results for both normal and hypervalent phosphorus compounds,[31] and addition in the NDDO approximation would increase the number of two-electron integrals to be evaluated five fold,[31] which would slow the calculations to an unacceptable level, a different way of parametrically adding the flexibility required to calculate both normal and hypervalent compounds was sought.

Variable Core-Core Repulsion Parameter, α

To determine where in the MNDO algorithm added flexibility might be the most productive, two molecular basis sets, one consisting of only trivalent phosphorus (P^{III}) compounds and the other of only pentavalent phosphorus (P^{V}) compounds, were used for independent parameterization trials. Starting with a set of values obtained by extrapolation of the parameter values of other elements, the normally optimized seven atomic parameters (U_{ss}, U_{pp}, β_s, β_p, Z_s, Z_p, α) were partially optimized according to the general procedure discussed previously[2] but using an improved algorithm.[39] Experimental data used for the optimization included heats of formation, dipole moments, ionization potentials and molecular geometries (Tables 1 and 2). The U_{ss} and U_{pp} parameters were relatively insensitive to the basis sets, changing only slightly from initial values (see below). All of the other parameters

	Initial	P^{III}	P^{V}
U_{ss}	-52.00	-52.50	-52.00
U_{pp}	-43.50	-43.09	-44.06
β_s	-8.40	-4.12	-9.14
β_p	-3.40	-6.95	-3.40
Z_s	2.30	2.45	2.30
Z_p	1.55	1.70	1.57
α	2.27	2.34	2.52

differed considerably more between P^{III} and P^{V} values. The effects of the differences in the bonding parameters between P^{III} and P^{V} are twofold. First, s-orbital participation in the bonding of P^{III} compounds was limited by the lower β_s and the higher β_p for P^{III} as compared to P^{V}. This is consistent with bond angles close to 90° for P^{III} compounds (Table 2). Secondly, overlap, and, hence, bonding interactions, were increased for P^{V} compounds by the lower Z_s and Z_p for P^{V}. This partially compensated for increased repulsion due to more atoms around phosphorus. The only purpose served by the change in α was to decrease the core-core repulsion in P^{V} compounds. The major problem with the current MNDO parameters lies in predicting the heats of formation of hypervalent compounds, which, as mentioned previously, are predicted far too positive. The geometries of P^{III} and P^{V} compounds, on the otherhand, are predicted fairly well.[28] Consequently, it appeared a variable α parameter

might provide the extra flexibility required for hypervalent molecules.

One possibility was to make α a function of the charge on phosphorus, similar to the functional dependence of β_d used for sulfur.[30] Accordingly, α was optimized for each of a series of molecules while holding the other six parameters fixed, and a linear relationship was looked for between the optimized α values and the corresponding charges on the phosphorus atom. However, there was no correlation.

Since the need to add flexibility was caused by the difference in the number of atoms bonded to phosphorus in P^{III} and P^{V} compounds, a correlation was looked for between the optimized α values and the sum of density matrix terms between the atomic orbitals on the phosphorus atom and all the orbitals on all other atoms. Again the correlation was poor, atoms not directly bonded to phosphorus adding too much to the sum. When the density matrix terms were multiplied by the corresponding fock matrix terms to dampen out contributions from atoms distal to the phosphorus, however, a reasonable correlation existed. Although the correlation coefficient was only 0.63 for data on 16 molecules (Table 3) the desired trend was evident, and the data did not take into account the effect of optimizing all parameters with the new, more flexible core-core repulsion term.

Based on these results, the core-core repulsion parameter for phosphorus, α_p, was made a function of density and fock matrix terms:

$$\alpha_p = APF + BPF \sum_B \sum_{\mu}^{P} \sum_{\upsilon}^{B} |P_{\mu\upsilon} F_{\mu\upsilon}| \qquad (1)$$

where μ was an atomic orbital on the phosphorus atom, υ was an atomic orbital on any other atom B, and $P_{\mu\upsilon}$ and $F_{\mu\upsilon}$ were the density and fock matrix terms between μ and υ. APF and BPF were new parameters to be optimized.

An undesirable side effect of this change soon became obvious. For geometry optimization, derivatives of the energy with respect to geometrical parameters, $\{q_i\}$, must be calculated:

$$\frac{dE_{TOT}}{dq_i} = \frac{\partial E_{TOT}}{\partial q_i} + \sum\sum_{\mu\upsilon} \frac{\partial E_{TOT}}{\partial P_{\mu\upsilon}} \frac{\partial P_{\mu\upsilon}}{\partial q_i} \qquad (2)$$

where E_{TOT} is the sum of the electronic energy, E_{el}, and the core-core repulsion energy, E_c, and $P_{\mu\upsilon}$ has it previous meaning. Equation (2) can be broken down into its components:

$$\frac{dE_{el}}{dq_i} = \frac{\partial E_{el}}{\partial q_i} + \sum\sum_{\mu\upsilon} \frac{\partial E_{el}}{\partial P_{\mu\upsilon}} \frac{\partial P_{\mu\upsilon}}{\partial q_i} \qquad (3)$$

$$\frac{dE_c}{dq_i} = \frac{\partial E_c}{\partial q_i} + \sum\sum_{\mu\upsilon} \frac{\partial E_c}{\partial P_{\mu\upsilon}} \frac{\partial P_{\mu\upsilon}}{\partial q_i} \qquad (4)$$

For variationally optimized electronic wavefunctions, the energy is first order invariant to small changes in the density matrix (since the energy is minimized with respect to the wave function in an SCF cycle), and all terms in the double summation on the right hand side (RHS) of equation (3) are zero. E_c normally has no functional dependence on the density matrix, so the terms in the double summation on the RHS of equation (4) are zero, also. Derivatives can, therefore, be calculated with a fixed density matrix.

The functional variation of α_p changed this, however. While the electronic wavefunction remained variationally optimized, the dependence of α_p on the density matrix prevented the terms in the double summation in equation (4) from being necessarily zero. Hence, SCF calculations had to be used to calculate the derivatives of E_c. This made geometry calculations much slower. This was already the case for some calculations, e.g. half-electron[40] calculations of radicals, though, and the problem was not compounded by an expanded basis set, as it was with sulfur.[30]

Several sets of parameter values were used as initial points for optimization of the, now, eight atomic parameters (U_{ss}, U_{pp}, β_s, β_p, Z_s, Z_p, APF, BPF) using the functional dependence of α_p shown in equation (1). The molecules used in the basis set were also varied. As previously indicated,[31] the region of the parameter hypersurface in which the parametrization became fixed was strongly dependent on these selections, as well as on the weights assigned to the various

observables used for the parameterization. While, for the first time, some P^{III} and P^{V} compounds were calculated relatively accurately with the same set of parameters, totally satisfactory results were never obtained.

A major problem area involved compounds with multiple bonds to phosphorus, such as P_2 and PN. Optimized α values for these compounds were surprisingly high, similar to values required for P^{V} compounds, while the sum of the product of density and fock matrix terms were far too small to give high values for α_p. These molecules are not that significant, and it would have been no great loss if the algorithm for α_p worked for all molecules except them. But these results did raise an interesting question: In what way are P_2 and PN more similar to P^{V} compound than P^{III} compounds? The answer lies in the amount of s-orbital participation in bonding. Assuming a bonding picture for P_2 and PN similar to that for N_2, i.e. a σ bond of sp hybrids and two P_π-P_π bonds, s-orbital participation in bonding is significant.[41] Similarly, s-orbital participation in bonding of P^{V} compounds is also significant, while that in normal P^{III} compounds is slight.

Given these considerations, a linear correlation was looked for between the optimized α values used previously and the ratio of interactions with the phosphorus s-orbital to those with the phosphorus p-orbitals, using the same type of summation as before for measuring interactions:

$$\frac{\sum_{B}\sum_{\upsilon}^{B}|P_{\mu\upsilon}F_{\mu\upsilon}|}{\sum_{B}\sum_{\mu'}^{P}\sum_{\upsilon}^{B}|P_{\mu'\upsilon}F_{\mu'\upsilon}|} \qquad (5)$$

where μ is now the s-orbital and μ' is a p-orbital on the phosphorus atom. For this data (Table 3), the correlation was significantly better (correlation coefficient = 0.77), so the functional dependence of α_p was changed accordingly:

$$\alpha_p = APF + BPF\left(\frac{\sum_{B}\sum_{\upsilon}^{B}|P_{\mu\upsilon}F_{\mu\upsilon}|}{\sum_{B}\sum_{\mu'}^{P}\sum_{\upsilon}^{B}|P_{\mu'\upsilon}F_{\mu'\upsilon}|}\right) \qquad (6)$$

Optimization using equation (6) for α_p did give better results. The average absolute error in the heats of formation for P^{III} and P^{V} compounds combined dropped to below 20 kcal/mol. However, another problem surfaced. The planar form of PH_3 was calculated to be more stable than the pyramidal form. By including planar PH_3 in the basis set, with a heat of formation derived from the known 37 kcal/mol barrier to inversion for PH_3,[42] and optimizing one-center electron-electron repulsion integral values (g_{ss}, g_{sp}, g_{pp}, $g_{pp'}$, h_{sp}) along with the other parameters, the average error in the heats of formation again dropped below 20 kcal/mol. Calculated dipoles, ionization potentials and molecular geometries were also satisfactory. Still, inversion barriers remained a problem. While the calculated inversion barrier for PH_3 was a

reasonable 25 kcal/mol, planar $P(CH_3)_3$ was more stable than pyramidal $P(CH_3)_3$ at the same time the inversion barrier for PF_3 was calculated unreasonably high (150 kcal/mol). Various attempts were made to obtain a set of parameters giving more consistent results, including changing to the AM1 hamiltonian with the hope that the better core-core repulsion functions for groups bonded to phosphorus would help lead to the consistency desired. All of these attempts failed.

Ultimately, the cause of the difficulty became apparent. The variable α_p function was designed to give larger α_p values (and, hence, smaller core-core repulsion energies) with greater s-orbital participation in bonding. This was expected to correct problems with P^V compounds and compounds involving multiple bonds to phosphorus. Not considered was the fact that the planar transition state for the inversion of pyramidal P^{III} compounds also has significant s-orbital participation in bonding. As a result, the core-core repulsion energy in these planar species tended to be underestimated. Trying to compensate for this in the optimization of the parameters led to the inconsistent results. As this deficiency goes to the very root of the logic behind the functional dependence of α_p, this approach was abandoned.

New Parameterization Algorithm

At the same time the fatal deficiency in the calculations involving the variable α parameter for phosphorus was becoming evident, a new parameterization algorithm[43] which was expected to be more effective than the previously used method, became available. The older method,[39] which was used for all recent parameterization, did not involve the calculation of full second derivatives for determining the search direction during optimization. As a result, the search directions determined in the program were very tentative.[44] Parameterization was typically started by allowing only one parameter to vary at a time. As the optimization began to stabilize, more than one parameter would be allowed to change simultaneously. Movement around the parameter surface was relatively slow, and could be influenced by the researcher. The decision as to when to call the parameterization complete was subjective, dependent upon attaining results which gave optimal agreement with experimental data *and* with chemical intuition.

In contrast, the new parameterization included the calculation of full second derivatives, and could, therefore, "see" the parameter surface more fully. As a result, a more powerful optimization routine could be employed than in the previous program, and the program was designed to work more autonomously. All parameters set to optimize were varied simultaneously, and the program found the best mathematical solution (the global minimum)

based on the experimental data used as input.[44]

When the new parameterization method was used to optimize all 12 MNDO parameters for P^{III} and P^{V} combined, including the one-center, two-electron integrals, results were, at first, encouraging. The program clearly reached areas of the parameter surface which the previously used method would never have reached had it not started in the general vicinity.[45] The average absolute error in the heats of formation for molecules in the basis set typically would decrease to below 15 kcal/mol within at least 20 cycles of the parameterization (a one-night run), even when the average error for the initial set of parameters was over 80 kcal/mol. Improvements in ionization potentials, dipoles and geometries were similar. If the same molecular basis set was used a second time but with a significantly different initial set of parameter values, the program would converge to essentially the same set of values as previously, indicating it was the global minimum.

Results were, however, entirely basis set dependent. Agreement with any experimental data not used in obtaining a given set of parameter values was often extremely poor. For example, while the average error in the heat of formation of molecules in the basis set might be under 12 kcal/mol, the inversion barrier for PH_3 could be overestimated by 90 kcal/mol and the proton affinity underestimated by 50 kcal/mol. If PH_4^+ was then added to the basis set, the set of parameter values obtained might again give good

results for the basis set molecules, including PH_4^+, but the proton affinity for $P(CH_3)_3^+$ could be underestimated by 40 kcal/mol.

The strength of the new parameterization was, at once, also a weakness. It was able to move around the parameter surface so freely, it appeared to curve fit too specifically to the basis set. Despite many different approaches to using the program, such as restricting different parameters from optimizing during alternate runs (to try to stabilize in a specific region of the surface) or optimizing only P^{III} or P^{V}, it was not possible to influence the parameterization; once a basis set was selected and weights assigned to the observables, the program was in complete control. In the end, no satisfactory set of parameter values was obtained.

Conclusions

The primary shortcoming of MNDO calculations for hypervalent compounds is the inability of calculated bonding interactions to compensate for the increased repulsion due to a larger number of atoms around the hypervalent atom than in normal valency compounds. While MNDO has been shown to handle pentavalent adducts of silicon,[46] effectively reproducing the three-center four-electron bonds thought to best describe the primary bonding of hypervalent species,[32,33,34] there is no assurance the stabilization derived from forming the adduct is fully estimated by the MNDO calculations. The problem is even more difficult for phosphorus than for silicon.

While all silicon valence atomic orbitals are involved in bonding in normal valency compounds and in the pentavalent adducts, this is not true for phosphorus. Consequently, parameter values must ideally restrict s-orbital participation in the bonding of pyramidal P^{III} compounds while allowing significant bonding participation in P^{V} compounds and compounds involving multiple bonds to phosphorus. Unfortunately, this flexibility is difficult to produce in the MNDO and AM1 formalisms.

To compensate for these factors, the core-core repulsion parameter was modified to decrease the core-core repulsion energy as bonding interactions with phosphorus increase. The difficulty, however, was finding a suitable measure of bonding interactions. The two functions of density and fock matrix terms tried were ultimately found to have fatal flaws. It is difficult to see any other functional dependence that would have a better chance of success.

The best solution still appears to be the addition of d-AOs. While previous attempts to add d-AOs for phosphorus[31] and sulfur[30] failed to give satisfactory results, the CNDO approximation under which they were employed was no doubt a major part of the problem. *Ab initio* results indicate exchange terms are very important in the involvement of d-AOs.[33] Consequently, a better approximation will probably be necessary to incorporate d-orbitals effectively.

While efforts to parameterize for phosphorus using a new, "more powerful" parameterization algorithm failed to yield a useful set of parameter values, they did serve to point out the essentiality of a chemist's input into the parameterization. As pointed out earlier,[31] rather than being a smooth surface with one "true" minimum, the parameter surface no doubt has many minima. While they may all be stationary points (i.e. mathematical solutions), they will not all necessarily correspond to solutions which will make chemical sense. Consequently, the researcher must have the ability to influence the parameterization. When a program becomes so powerful that this ability is restricted, there is a good chance that the solution obtained, though mathematically correct, may not be chemically useful.

Table 1. Experimental heats of formation, dipole moments and ionization potentials of phosphorus compounds.

Molecule	Heat of Formation[a]	Dipole Moments[b]	Ionization Potentials[c]
P_2	34.5[d]		10.62[f]
P_4	14.1[d]		
PH_3 (C_{3v})	1.3[d]	0.57[e]	10.59[f]
P_2H_4	5.0[d]	0.92[g]	9.69[f]
CH_3PH_2		1.10[h]	9.6[f]
$(CH_3)_2PH$		1.23[i]	9.1[f]
$(CH_3)_3P$	24.2[j]	1.19[h]	8.6[f]
$C_6H_5P(CH_3)_2$			8.45[f]
C_2H_4PH[k]		1.12[l]	
$C_3H_5PH_2$[m]		1.16[l]	
C_5H_5P[n]			9.2[f]
PCH		0.39[i]	
$PCCH_3$			9.89[f]
PN	26.3[d]	2.75[l]	
P_4O_6	-378[o]		10.55[f]
$P(OCH_3)_3$	-168.6[i]		
PHF_2		1.35[i]	11.0[f]
PF_2	-219.6[d]	1.03[i]	12.2[f]
H_2PPF_2		1.71[i]	
PH_2CF_3		1.92[l]	11.2[f]
PF_2CH_3		2.06[l]	10.35[f]
PF_2CN			11.9[p]
PF_2NH_2		2.58[l]	10.9[p]
PF_2Cl		0.89[l]	11.5[p]
PCl_2CH_3			9.85[f]
$PCl_2C_2H_5$	-74.4[j]		9.70[f]
PCl_3	-68.6[d]	0.58[l]	10.5[f]
PBr_3	-33.3[d]		10.0[f]
PI_3			9.15[f]
$(CH_3)_3PCH_2$			6.8[f]
$(CH_3)_3PNC_2H_5$	-22.7[j]		
$(CH_3)_3PNH$			8.2[f]
$(CH_3)_3PO$	-103.8[j]		9.9[f]
$(CH_3)(OH)_2PO$	-240.2[j]		
$(C_2H_5)(OH)_2PO$	-241.3[j]		
$(OCH_3)_3PO$			10.8[f]
$(OC_2H_5)_3PO$	-283.6[j]		
HF_2PO		2.65[l]	

Table 1. (Contd.)

Molecule	Heat of Formation[a]	Dipole Moments[b]	Ionization Potentials[c]
F_3PO	-289.5[d]	1.87[e]	13.5[p]
CH_3Cl_2PO	-132.9[j]		
Cl_3PO	-133.5[d]	2.54[l]	
Br_3PO	-94.0[q]		11.0[f]
PF_3H		1.32[l]	
PF_4	-381.4[d]		15.5[f]
PCl_5	-89.6[d]		10.7[f]
P_4O_{10}	-672[o]		

[a] kcal/mol. [b] debye. [c] Vertical ionization potential in electron volts (eV). [d] Ref. 47. [e] Ref. 48. [f] Ref. 49. [g] Ref. 50. [h] Ref. 51. [i] Ref. 52. [j] Ref. 53. [k] Phosphirane. [l] Ref. 54. [m] Cyclopropyl phosphine. [n] Phosphabenzene. [o] Ref. 55. [p] Ref. 56. [q] Ref. 57.

Table 2. Experimental geometrical parameters of phosphorus compounds.[a]

Molecule	Geometrical Parameters	Reference
P_2	PP 1.894	58
P_4	PP 2.21	42
PH	PH 1.433	58
PH_2	PH 1.418; HPH 91.7	58
PH_3	PH 1.420; HPH 93.3	58
P_2H_4	PP 2.219; PH^1 1.414; PH^2 1.417; H^1PH^2 92.0; H^1PP 94.3; H^1_2PP 99.1; ϕ[b] 74.0	58
CH_3PH_2	CP 1.863; PH 1.414; CH 1.093; CPH 97.5; HPH 93.4; HCH 109.7; methyl group tilted away from PH_2 moiety 2.00	58
$P(CH_3)_3$	CP 1.843; CH_s 1.112; CH_a 1.090; PCH 111.4; PCH_s 109.8; HCH_s 108.2; HCH_a 109.4; CPC 98.9 (s = in plane, a = out of plane)	58
C_2H_4PH[c]	CP 1.867; PH 1.428; CC 1.502; CH_{cis} 1.092; CH_{trans} 1.093; CPH 95.2; CCH_{cis} 118.0; CCH 117.5; HCH 114.4	58
CH_2PH	CP 1.673; PH 1.420; CH_{cis} 1.09; CH_{trans} 1.09; $H_{cis}CP$ 124.4; $H_{trans}CP$ 118.4; CPH 97.4	59
C_5H_5P[d]	$CP_{1,3}$ 1.733; $C^2C^3_{1,4}$ 1.413; $C^3C^4_{1,3}$ 1.384; $C^2PC^6_{4,5}$ 101.1; $PC^2C^3_{2,3}$ 124.4; $C^2C^3C^4_{2,3,4}$ 123.7; $C^3C^4C^5$ 122.8; PC^2H 118.2; C^2C^3H 114.7	60
CP	CP 1.56	61
PCH	CP 1.542; CH 1.068; PCH 180	58
PN	PN 1.491	58
PO	PO 1.474	58
PO_2	PO 1.467; OPO 135.3	62
P_4O_6	PO 1.638; POP 126.4; OPO 99.8	60

Table 2. (Contd.)

Molecule	Geometrical Parameters	Reference
PHF_2	PH 1.412; PF 1.582; HPF 96.3; FPF 99.0	58
PF_3	PF 1.563; FPF 97.7	58
PF_2CN	CP 1.811; PF 1.567; CN 1.158; FPF 99.1; FPC 97.1; PCN 171.5	58
PF_2OCH_3	PF 1.591; PO 1.560; CO 1.446; CH 1.090; FPF 94.8; OPF 102.2; COP 123.7; HCH 110.5	58
PF_2NH_2	PF 1.587; PN 1.650; NH 1.002; NH 0.981; FPF 94.6; FPN 100.6; PNH_{cis} 123.1; PNH_{trans} 119.7; HNH 117.2	58
PF_2Cl	PF 1.571; CCl 2.030; FPF 97.3; FPCl 99.2	58
PCl_3	PCl 2.043; ClPCl 100.1	58
PBr_3	PBr 2.220; BrPBr 101.0	60
PI_3	PI 2.46; IPI 100	63
PS	PS 1.92	61
HF_2PO	PH 1.387; PF 1.539; PO 1.437; HPO 117.9; FPO 116.3; FPF 99.8	58
F_3PO	PO 1.437; PF 1.522; FPF 101.14	58
Cl_3PO	PO 1.455; PCl 1.989; ClPCl 103.7	58
PF_5	PF_{ax} 1.576; PF_{eq} 1.529	64
PCl_2F_3	PF_q 1.539; PF_{ax} 1.591; PCl 2.001; $F_{ax}PF_{eq}$ 89.3; ClPCl 122.2	65
PCl_5	PCl_{ax} 2.124; PCl_{eq} 2.020	60
P_4O_{10}	PO^1 1.429; PO^2 1.604; PO^2P 123.5; O^2PO^2 101.6; O^2PO^1 116.5	60
HF_2PS	PH 1.392; PF 1.551; PS 1.867; SPF 117.4; SPH 119.2; SPH 119.2; FPF 98.6	58

Table 2. (Contd.)

Molecule	Geometrical Parameters	Reference
F_3PS	PS 1.87; PF 1.53; FPF 100.3	58

[a]Bond lengths are in angstroms. Bond angles and dihedral angles are in degrees. [b]Dihedral measured from cis position. Hydrogens designated H^2 are the nearly eclipsed pair. [c]Phosphirane. [d]Phosphabenzene.

Table 3. Optimized α values and corresponding values for functions of the density and fock matrix terms.[a]

Molecule	α	$\sum_{B}\sum_{\mu}^{P}\sum_{\nu}^{B}\lvert P_{\mu\nu} F_{\mu\nu}\rvert^{b}$	$\frac{\sum_{B}\sum_{\mu}^{B}\lvert P_{\mu\nu} F_{\mu\nu}\rvert^{b}}{\sum_{B}\sum_{\mu'}^{P}\sum_{\nu}^{B}\lvert P_{\mu'\nu} F}$
PCl_3	2.254	22.54	0.0602
$PCl_2C_2H_5$	2.309	24.50	0.0903
PBr_3	2.330	21.49	0.0602
P_4O_6	2.340	32.50	0.1107
P_2H_4	2.346	23.79	0.1259
P_4	2.364	18.67	0.0501
PH_3	2.370	25.00	0.1266
PF_3	2.392	37.27	0.1066
$(CH_3)_3PNC_2H_5$	2.423	38.96	0.2671
PN	2.450	25.96	0.1258
P_4O_{10}	2.461	49.74	0.3018
Cl_3PO	2.481	40.18	0.3055
PCl_5	2.487	35.14	0.2982
F_3PO	2.488	54.84	0.2914
$(CH_3)_3PO$	2.513	43.07	0.2896
PF_5	2.526	59.87	0.2859

[a] α values were optimized for the optimized geometries using experimental values for heats of formation, dipole moments and ionization potentials given in Table 1. The values for other parameters held fixed were: U_{ss} = -52.50, U_{pp} = -43.07, β_s = -4.12, β_p = -6.95, Z_s = 2.45, Z_p = 1.70. [b] See text for explanation of terms.

References

1. Bingham, R. C.; Dewar, M. J. S.; Lo, D. H. *J. Am. Chem. Soc.* **1975**, *97*, 1285, 1294, 1302.

2. Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899, 4907.

3. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

4. (a) Pople, J. A.; Santry, D. P.; Segal, G. A. *J. Chem. Phys.* **1965**, *43*, S129. (b) Pople, J. A.; Beveridge, D. L.; Dobosh, P. A. *J. Chem. Phys.* **1967**, *47*, 2026. (c) Pople, J. A.; Beveridge, D. L. Approximate Molecular Orbital Theory, McGraw Hill Book Company: New York, 1970.

5. MINDO/3 is based on the INDO approximation.

6. MNDO and AM1 are based on the NDDO approximation.

7. (a) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69. (b) Hall, G. G. *Proc. R. Soc. London, Ser. A*, **1951**, *205*, 541.

8. (a) Hartree, D. R. *Proc. Cambridge Phil. Soc.* **1928**, *24*, 89. (b) Fock, V. *Z. Physik* **1930**, *61*, 126.

9. In the HF/SCF method, each electron is treated as moving in the average field of all the other electrons. As such the instantaneous adjustment of one electron to the movement of another electron is not taken into account.

10. Dewar, M. J. S. *J. Phys. Chem.* 1985, *89*, 2145.

11. Dewar, M. J. S.; Rzepa, H. S. *J. Am. Chem. Soc.* 1978, *100*, 784.

12. (a) MINDO/3: Novoa, J. J. *J. Mol. Struct. Theochem* 1986 *136*, 361. (b) MNDO: Olivella, S.; Urpi, F.; Vilarrasa, J. *J. Comput. Chem.* 1984, *5*, 230. Chandrasekhar, J.; Andrade, J. G.; Schleyer, P. v. R. *J. Am. Chem. Soc.* 1981, *103*, 5609. (c) AM1: Chapter 3 of this dissertation.

13. Dewar, M. J. S.; Storch, D. M. *J. Am. Chem. Soc.* 1985, *107*, 3898.

14. Dewar, M. J. S.; Rzepa, H. S. *J. Am. Chem. Soc.* 1978, *100*, 777.

15. Dewar, M. J. S.; McKee, M. L. *J. Am. Chem. Soc.* 1977, *99*, 5231.

16. (a) MINDO/3: Bingham, R. C.; Dewar, M. J. S.; Lo, D. H. *J. Am. Chem. Soc.* 1975, *97*, 1307. (b) MNDO: Dewar, M. J. S.; Rzepa, H. S. *J. Am. Chem. Soc.* 1978, *100*, 58. (c) AM1: Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F. Stewart, J. J. P.

To be published.

17. MNDO: Davis, L. P.; Guidry, R. M.; Williams, J. R.; Dewar, M. J. S.; Rzepa, H. S. *J. Comput. Chem.* 1981, *2*, 433.

18. (a) MINDO/3: Dewar, M. J. S.; Lo, D. H.; Ramsden, C. A. *J. Am. Chem. Soc.* 1975, *97*, 1311. (b) MNDO: Dewar, M. J. S.; Healy, E. F.; Stewart, J. J. P; Grady, G. L. *Organometallics* 1986, *5*, 375.

19. (a) MINDO/3: Ref. 18(a). (b) MNDO: Dewar, M. J. S.; McKee, M. L.; Rzepa, H. S. *J. Am. Chem. Soc.* 1978, *100*, 3607.

20.(a) MINDO/3: Ref. 18(a) (b) MNDO: Dewar, M. J. S.; Reynolds, C. H. *J. Comput. Chem.* 1986, *7*, 140.

21. (a) MINDO/3: Ref. 16(a). (b) MNDO: Ref. 19(b). Dewar, M. J. S.; Rzepa, H. S. *J. Comput. Chem.* 1983, *4*, 158. (c) AM1: Ref. 16(c).

22. (a) MNDO: Dewar, M. J. S. ; Healy, E. F.; *J. Comput. Chem.* 1983, *4*, 542. (b) AM1: Ref. 16(c).

23. (a) MNDO: Dewar, M. J. S.; Grady, G. L.; Stewart, J. J. P. *J. Am. Chem. Soc.* 1984, *106*, 6771. (b) AM1: Dewar, M. J. S.; Kuhn, D. R.; Healy, E. F.; Stewart, J. J. P. To be published.

24. (a) MNDO: Dewar, M. J. S.; Healy, E. F.; Stewart, J. J. P. *J. Comput. Chem.* **1984**, *5*, 358. (b) AM1: Ref. 16(c).

25. MNDO: Dewar, M. J. S.; Grady, G. L.; Merz, K.M., Jr.; Stewart, J. J. P. *Organometallics* **1985**, *4*, 1964.

26. MNDO: Dewar, M. J. S.; Holloway, M. K.; Grady, G. L.; Stewart, J. J. P. *Organometallics* **1985**, *4*, 1973.

27. The average error is based on values calculated by McKee[28] and the experimental values given in Table 1, this work. This error for P^{III} is much larger than originally thought by McKee due to several experimental values used here being different from those used by him.

28. McKee, M. L. Unpublished results.

29. (a) Pople, J. A.; Segal, G. A. *J. Chem. Phys.* **1965**, *43*, 5136. (b) Ref. 4(c).

30. (a) Reynolds, C. H. Ph.D. Dissertation, The University of Texas at Austin, 1984. (b) Ref. 20 (b).

31. Storch, D. M. Ph.D. Dissertation, The University of Texas at Austin, 1985.

32. Schmidt, M. W.; Yabushita, S.; Gordon, M. S. *J. Phys. Chem.* **1984**, *88*, 382.

33. Magnusson, E. *Int. Rev. Phys. Chem.* **1986**, *5*, 147.

34. Kutzelnigg, W. *Angew. Chem., Int. Ed. Engl.* **1984**, *23*, 272.

35. Raghavachari, K.; Haddon, R. C.; Binkley, J. S. *Chem. Phys. Lett.* **1985**, *122*, 219.

36. Trinquier, G.; Daudey, J.-P.; Komiha, N. *J. Am. Chem. Soc.* **1985** *107*, 7211.

37. Ahlrichs, R.; Brode, S.; Ehrhardt, C. *J. Am. Chem. Soc.* **1985**, *107*, 7260.

38. Haleri, E. A.; Bock, H.; Roth, B. *Inorg. Chem.* **1984**, *23*, 4376.

39. Stewart, J. J. P. Unpublished work. For a discussion of the method, see Healy, E. F.; Ph.D. Dissertation, The Univerity of Texas at Austin, 1984.

40. (a) Dewar, M. J. S.; Hashmall, J. A.; Venier, C. G. *J. Am. Chem. Soc.* **1968**, *90*, 1953. (b) Dewar, M. J. S.; Trinajstic, J. *J. Chem. Soc. D:* **1970**, 646. (c) Dewar, M. J. S.; Trinajstic, N. *J. Chem. Soc. A:* **1971**, 1220.

41. *Ab initio* results again predict the sp hybridization model overestimates s-orbital participation in the bonding in P_2.[34]

42. Cotton, F.A.; Wilkinson, G. *Advanced Inorganic Chemistry, 4th Ed.*; John Wiley & Sons: New York, 1980, p. 443.

43. Stewart, J. J. P. Unpublished work.

44. Stewart, J. J. P. Personal communication.

45. Starting with U_{ss} = -52 eV and U_{pp} = -42 eV, the older parameterization routine tended to stabilize with a U_{ss} of approximately -48 to -54 eV and a U_{pp} of approximately -40 to -44 eV. In contrast, the newer routine tended to make U_{ss} and U_{pp} more equal, at approximately -40 eV. Experience has shown that the older program would likely never have reached this same region of the surface.

46. Dewar, M. J. S.; Healy, E. F. *Organometallics* 1982, *1*, 1705.

47. Wagman, D.D.; Evans, W. H.; Parker, V. B.; Schumm, R. H.; Halow, I.; Bailer, S. M.: Churney, K. L.; Nuttall, R. L. "The NBS Tables of Chemical Thermodynamic Properties: Selected Values for Inorganic and C_1 and C_2 Organic Substances in S.I. Units" *J. Phys. Chem. Ref. Data* 1982, *11*, Suppl. 2.

48. Demaison, J.; Dubrulle, A.; Hüttner. W.; Tiemann, E. *Numerical Data and Functional Relationships in Science and Technology, New Series, Group II*; Hellwege, K.-H.; Hellwege. A.M., Ed.; Springer-Verlag; Heidelberg, 1982; Vol. 14, Subvol.

a.

49. Levin, R.D.; Lias, S. G. *Ionization Potential and Appearance Potential Measurements, 1971-1981* NSRDS-NBS71, 1982.

50. Durig, J. R.; Carreira, L. A.; Odom, J. D. *J. Am. Chem. Soc.* 1974, *96*, 2688.

51. McClellan, A.L. *Tables of Experimental Dipole Moments* W. H. Freeman and Company: San Francisco, 1963.

52. McClellan, A.L. *Tables of Experimental Dipole Moments* Rahara Enterprises: El Cerrito, CA, 1974; Vol. 2.

53. Pedley, J. B.; Rylance, G. *Sussex-N.P.L. Computer Analysed Thermochemical Data: Organic and Organometallic Compounds*, Sussex University, 1977.

54. Demaison, J.; Hüttner, W.; Starck, B.; Buck, I.; Tischer, R.; Winnewisser, M. *Numerical Data and Functional Relationships in Science and Technology, New Series, Group II*; Hellwege, K.-H.; Hellwege, A.M., Ed.; Springer-Verlag: Heidelberg, 1974, Vol. 6.

55. Meunow, D. W.; Uy, O. M.; Margrave, J.L. *J. Inorg. Nucl. Chem.* 1970, *32*, 3459.

56. Siegbahn, K.; Allison, D. A.; Allison, J. H. *Handbook of Spectroscopy*, Robinson, J. W., Ed.; CRC Press:Cleveland, OH, 1974.

57. Hartley, S. B.; Holmes, W. S.; Jacques, J. K.; Mok, M. F.; McCoubrey, J. C. *Q. Rev. Chem. Soc.* **1963**, *17* 204.

58. Harmony, M. D.; Laurie, V. W.; Kuczkowski, R. L.; Schwendeman, R. H.; Ramsay, D. A.; Lovas, F. J.; Lafferty, W. J.; Maki, A. G. *J. Phys. Chem. Ref. Data* **1979**, *8*, 619.

59. Kroto, H. W.; Nixon, J. F.; Ohno, K. *J. Mol. Spectrosc.*, **1981**, *90*, 367.

60. Callomon, J. H.; Hirota, E.; Kuchitsu, K.; Lafferty, W. J.; Maki, A. G.; Pote, C. S.; Buck, I.; Starck, B. *Numerical Data and Functional Relationships in Science and Technology, New Series, Group II*; Hellwege, K.-H.; Hellwege, A.M., Ed.; Springer-Verlag: Heidelberg, 1976; Vol. 7.

61. Stull, D. R.; Prophet, J. *JANAF Thermochemical Tables*. NSRDS-NBS37, 1971.

62. Kawaguchi, K.; Saito, S.; Hirota, E.; Ohashi, N. *J. Chem. Phys.* **1985**, *82*, 4893.

63. *Tables of Interatomic Distances and Configurations in Molecules and Ions*, Mitchell, A. D.; Cross, L. C., General, Ed.; The Chemical Society Burlington House: London, 1958.

64. Kurimura, H.; Yamamoto, S.; Egawa, T.; Kuchitsu, K. *J. Mol. Struct.* **1986**, *140*, 79.

65. French, R. J.; Hedberg, K.; Shreeve, J.M.; Gupta, K. D. *Inorg. Chem.* **1985**, *24*, 2774.

Chapter 3

EVALUATION OF AM1 CALCULATED PROTON AFFINITIES AND DEPROTONATION ENTHALPIES

Introduction

Proton transfer reactions play a basic role in chemistry, in particular in biochemistry. A knowledge of the proton affinities (PA) of bases, and of the deprotonation enthalpies (DPE) of acids, is, therefore, essential. While major progress has been made in recent years in the development of experimental techniques for measuring PAs, these have necessarily been limited to the gas phase, and results are available only for a very limited number of ions and molecules[1]. If PAs could be calculated theoretically with sufficient accuracy by some quantum chemical procedure, this would be of major value, because calculations, if feasible, can be carried out much more quickly and at much less cost than experiments, and they are, of course, not limited by the physical properties of the species being studied.

In connections such as this, the only theoretical procedures that need to be considered are *ab initio* ones based on the Roothaan-Hall[2] SCF MO approximation and the semiempirical SCF MO methods (MINDO/3[3], MNDO[4], AM1[5]) developed here. Other alternatives

(e.g. CNDO/2) are too inaccurate and unreliable[6].

Numerous *ab initio* calculations of absolute and relative PAs and DPEs have been reported for compounds derived from carbon, hydrogen, nitrogen and oxygen[7-9]. Many of these can, however, be eliminated from consideration because of failure to optimize geometries. As would be expected, the results from the others depend greatly on the basis set used and on the allowance (if any) made for electron correlation. While the number of calculations is too large for the results to be compared here in detail, it seems clear that a reasonably large basis set must be used if the average errors in DPEs and PAs are to be kept within reasonable limits, e.g. 10 kcal/mol. In the case of anionic bases, it is also necessary to include diffuse AOs in the basis set[7]. Calculations at this level become very expensive for larger molecules, and the results are still by no means exact.

MNDO[4] is now a well established procedure for calculating molecular properties[10]. While it was parametrized to reproduce ground state properties of neutral closed-shell molecules, it also gives good results for radicals, carbenes, and positive and negative ions. The only exceptions are anions in which the charge is concentrated on a single atom, where the calculated heats of formation are much too positive[11]. This error is probably[11] due to failure to allow for orbital expansion accompanying large localized negative charges. It was, therefore, reasonable to hope that MNDO

might provide satisfactory estimates of DPEs and PAs, and two recent studies[7,11] have indeed shown that the results are comparable with those from *ab initio* procedures that require a thousand times more computing time. The errors were, however, larger than desirable, particularly in the case of certain molecules involving features known to present problems in MNDO. These problems have now been mostly overcome in a new "third generation" semiempirical model, AM1[5]. In this study, then, the performance of AM1 in calculations of DPEs and PAs is tested using calculation results for a large number of molecules for which apparently satisfactory experimental values are known.

Procedure

The calculations were carried out using the standard AM1[5] procedure, as implemented in the AMPAC package of computer programs[12]. All geometries were optimized by minimizing the energy with respect to all geometrical variables using the DFP method[13] incorporated in AMPAC and without making *any* assumptions. Various geometries of protonated methane and ethane were characterized as either true minima, transition states, or "hilltops" by diagonalizing their Hessian (force constant) matrices and looking for zero, one, or two or more, negative eigenvalues, respectively[14].

DPEs and PAs were found by difference from the calculated heats of formation of the parent molecule and the derived cation or anion. Since AM1 gives a very poor estimate of the heat of formation of H^+ (calcd. 314.9; obsd. 367.2[15] kcal/mol), the experimental value was used in calculating DPEs and PAs.

Results and Discussion

A. Proton Affinities

The PA of a compound (B) is defined as *minus* the heat of reaction for its combination with a proton;

$$B + H^+ \rightarrow HB^+$$

$$PA(B) = \Delta H_f(H^+) + \Delta H_f(B) - \Delta H_f(HB^+) \qquad (1)$$

Table 1 lists calculated and experimental PAs for 60 compounds together with errors ($\delta\Delta H_f(HB^+)$) in the heats of formation calculated for the ionic species, these being found by using Eq (2);

$$\delta\Delta H_f(HB^+) = \delta\Delta H_f(B) - \delta PA(B) \qquad (2)$$

where $\delta\Delta H_f(B)$ and $\delta PA(B)$ are, respectively, the errors in the heat of formation and in the PA of B. Included are 9 carbon bases (the highest calculated PA is for protonation of a carbon atom), 33 nitrogen bases, 12 oxygen bases, and 6 bifunctional bases discussed separately because of hydrogen bonding considerations.

Table 1. Comparison with Experiment of AM1 Heats of Formation and Proton Affinities (kcal/mol).

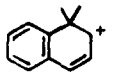
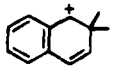
B	HB^+	Calcd ΔH_f B	Calcd ΔH_f HB^+	Proton Affinity Calcd	Proton Affinity Exptl[b]	Proton Affinity Error	Exptl ΔH_f(B)[d]	Error in Calcd ΔH_f B	Error in Calcd ΔH_f HB^+
Carbon Bases									
CH_4	CH_5^+	-8.8[a]	224.4	134.0	132.0	2.0	-17.8	9.0	7.0
					134.7[c]	-0.7			9.7
CH_3CH_3	$CH_3CH_4^+$	-17.4[a]	208.4	141.4	143.6	-2.2	-20.0	2.6	4.8
					146.9[c]	-5.5			8.1
$CH_2=CH_2$	$CH_3CH_2^+$	16.5[a]	216.8	166.9	162.6	4.3	12.5	4.0	-0.3
$CH_3CH=CH_2$	$CH_3\overset{+}{C}HCH_3$	6.6[a]	191.9	181.9	179.5	2.4	4.8	1.8	-0.6
	$CH_3CH_2\overset{+}{C}H_2$		211.7	162.1					
		22.0[a]	206.0	183.2	181.3	1.9	19.8	2.2	0.3
		40.6[a]	213.2	194.6	194.7	-0.1	35.9	4.7	4.8
			216.7	191.1					
		14.5[a]	193.7	188.0					
			197.0	184.7					
			192.6	189.1	189.8	-0.7	12.0	2.5	3.2
HC≡CH	$H_2C=\overset{+}{C}H$	54.8[a]	261.5	160.5	153.3	7.2	54.5	0.3	-6.9
$CH_3C\equiv CH$	$CH_3\overset{+}{C}=CH_2$	43.4[a]	233.7	176.9	182	-5	44.2	-0.8	4
	$CH_3CH=\overset{+}{C}H$		251.2	159.4					

Table 1. (Contd.)

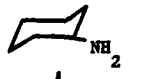
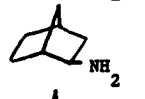
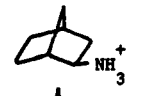
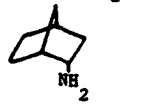
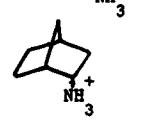
B	HB^+	Calcd ΔH_f B	Calcd ΔH_f HB^+	Proton Affinity Calcd	Proton Affinity Exptl[b]	Proton Affinity Error	Exptl $\Delta H_f(B)$[d]	Error in Calcd ΔH_f B	Error in Calcd ΔH_f HB^+
Nitrogen Bases									
NH_3	NH_4^+	-7.3[a]	150.6	209.3	204.0	5.3	-11.0[g]	3.7	-1.6
CH_3NH_2	$CH_3NH_3^+$	-7.4[a]	148.7	211.1	214.1	-3.0	-5.5	-1.9	1.1
$CH_3CH_2NH_2$	$CH_3CH_2NH_3^+$	-15.1[a]	138.7	213.4	217.0	-3.6	-11.3	-3.8	-0.2
$CH_3(CH_2)_2NH_2$	$CH_3(CH_2)_2NH_3^+$	-22.1[a]	130.8	214.3	217.9	-3.6	-16.8	-5.3	-1.7
$(CH_3)_2CHNH_2$	$(CH_3)_2CHNH_3^+$	-19.2[a]	131.1	216.9	218.6	-1.7	-20.0	0.8	2.5
$CH_3(CH_2)_3NH_2$	$CH_3(CH_2)_3NH_3^+$	-28.9	124.0	214.3	218.4	-4.1	-22.0	-6.9	-2.8
$(CH_3)_2CHCH_2NH_2$	$(CH_3)_2CHCH_2NH_3^+$	-25.2	125.2	216.8	218.8	-2.0	-23.6	-1.6	0.4
$(CH_3)_3CNH_2$	$(CH_3)_3CNH_3^+$	-21.2	125.6	220.4	220.8	-0.4	-28.9	7.7	8.1
$(CH_3)_3CCH_2NH_2$	$(CH_3)_3CCH_2NH_3^+$	-30.5	120.9	215.8	219.3	-3.5	-30[f]	-1	3
		-31.9	115.1	220.2	221.2	-1.0	-25.1	-6.8	-5.8
		-8.9	137.4	220.9	221.7	-0.8	-8[f]	-1	0
		-8.2	138.2	220.8	221.7	-0.9	-7[f]	-1	0
		20.7[a]	176.5	211.4	209.5	1.9	20.8	-0.1	-2.0
			181.5	206.4					
			203.5	184.4					

Table 1. (Contd.)

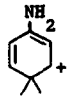
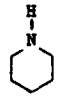
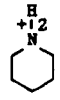
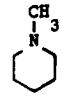
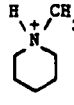
B	HB^+	Calcd ΔH_f B	Calcd ΔH_f HB^+	Proton Affinity Calcd	Proton Affinity Exptl[b]	Proton Affinity Error	Exptl[d] $\Delta H_f(B)$	Error in Calcd ΔH_f B	Error in Calcd ΔH_f HB^+
			178.9	209.0					
$(CH_3)_2NH$	$(CH_3)_2NH_2^+$	-5.6[a]	150.7	210.9	220.6	-9.7	-4.4	-1.2	8.5
$CH_3CH_2N(CH_3)H$	$CH_3CH_2\overset{+}{N}(CH_3)H_2$	-12.6	140.3	214.3	222.8	-8.5	-11[f]	-2	7
$(CH_3CH_2)_2NH$	$(CH_3CH_2)_2NH_2^+$	-17.8	131.5	217.9	225.9	-8.0	-17.3	-0.5	7.5
		-10.4	139.9	216.9	225.2	-8.3	-0.8	-9.6	-1.3
		-19.0	128.4	219.8	226.4	-6.6	-11.3	-7.7	-1.1
		-30.0	109.4	227.8	231.7	-3.9	-38[h]	8	12
$(CH_3)_3N$	$(CH_3)_3\overset{+}{N}H$	-1.7[a]	152.0	213.5	225.1	-11.6	-5.7	4.0	15.6
$(CH_3)_2NC_2H_5$	$(CH_3)_2\overset{+}{N}(C_2H_5)H$	-6.7	143.4	217.1	227.5	-10.4	-11[e]	4	14
$CH_3N(C_2H_5)_2$	$CH_3N(C_2H_5)_2H$	-11.6	135.0	220.6	230.0	-9.4	-17[e]	5	14
		-6.4	142.9	217.9	228.7	-10.8	-0.5[e]	-5.9	4.9
		-15.2	131.5	220.5	229.7	-9.2	-12[f]	-3	6

Table 1. (Contd.)

B	HB^+	Calcd ΔH_f B	Calcd ΔH_f HB^+	Proton Affinity Calcd	Proton Affinity Exptl[b]	Proton Affinity Error	Exptl $\Delta H_f(B)$[d]	Error in Calcd ΔH_f B	Error in Calcd ΔH_f HB^+
		-8.2	139.5	219.5	232.1	-12.6	-1.0	-7.2	5.4
		25.9	174.1	219.0	228.5	-9.5	37[f]	-11	-2
	$CH_3CHNH_2^+$	8.5	160.4	215.3	213.9	1.4	2[h]	7	6
		10.9		217.7					
		32.1[a]	184.2	215.1	220.8	-5.7	33.6	-1.5	4.2
		25.7	173.9	219.0	225.0	-6.0	23.7	2.0	8.0
		24.1	174.6	216.7	224.1	-7.4	25.4	-1.3	6.1
		24.2	173.3	218.1	225.2	-7.1	24.9	-0.7	6.4
HCN	$HCNH^+$	31.0[a]	214.8	183.4	171.4	12.0	32.3[e]	-1.3	-13.3
CH_3CH	CH_3CNH^+	19.3[a]	196.1	190.4	188.4	2.0	15.4	3.9	1.9

Table 1. (Contd.)

B	HB^+	Calcd ΔH_f B	Calcd ΔH_f HB^+	Proton Affinity Calcd	Proton Affinity Exptl[b]	Proton Affinity Error	Exptl $\Delta H_f(B)$[d]	Error in Calcd ΔH_f B	Error in Calcd ΔH_f HB^+
Oxygen Bases									
H_2O	H_3O^+	-59.2[a]	143.5	164.5	166.5	-2.0	-57.8[g]	-1.4	0.6
					166.4[c]	-1.9			0.5
CH_3OH	$CH_3OH_2^+$	-57.0[a]	138.3	171.9	181.9	-10.0	-48.2	-8.8	1.2
CH_3CH_2OH	$CH_3CH_2OH_2^+$	-62.7[a]	125.8	178.7	188.3	-9.6	-56.2	-6.5	3.1
$(CH_3)_3COH$	$(CH_3)_3COH_2^+$	-71.6[a]	107.8	187.8	193.7	-5.9	-74.7	3.1	9.0
$(CH_3)_2O$	$(CH_3)_2OH^+$	-53.2[a]	136.6	177.4	192.1	-14.7	-44.0	-9.2	5.5
$(C_2H_5)_2O$	$(C_2H_5)_2OH^+$	-64.4[a]	114.1	188.7	200.2	-11.5	-60.3	-4.1	7.4
$(CH_3)_3COCH_3$	$(CH_3)_3\overset{+}{C}O(CH_3)H$	-64.8	111.5	190.9	202.2	-11.3	-67.8	3.0	14.3
CO_2	CO_2H^+	-79.8[a]	147.8	139.6	130.9	8.7	-94.1[g]	14.3	5.6
					133.2[c]	6.4			7.9
H_2CO	H_2COH^+	-31.5[a]	161.3	174.4	171.7	2.7	-26.0[g]	-5.5	8.2
CH_3CHO		-41.6[a]	140.7	184.9	186.6	-1.7	-39.7	-1.9	-0.2
			142.3	183.3					
NH_2CHO		-44.7[a]	122.1	200.4	198.4	2.0	-44[h]	-1	-3
			122.7	199.8					

Table 1. (Contd.)

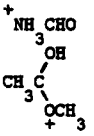
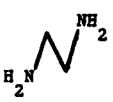
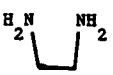
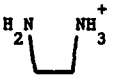
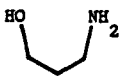
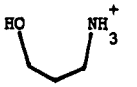
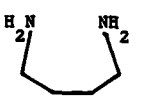
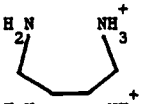
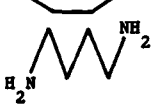
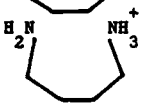
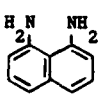
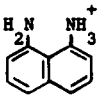
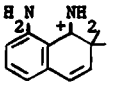
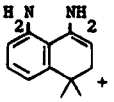
B	HB^+	Calcd ΔH_f B	Calcd ΔH_f HB^+	Proton Affinity Calcd	Proton Affinity Exptl[b]	Proton Affinity Error	Exptl $\Delta H_f(B)$[d]	Error in Calcd ΔH_f B	Error in Calcd ΔH_f HB^+
			132.5	190.0					
$CH_3CO_2CH_3$		-96.4[a]	76.0	194.8	197.8	-3.0	-98.4	2.0	5.0
Bifunctional Bases									
H_2NNH_2	$H_2NNH_3^+$	13.7[a]	184.0	196.9	204.7	-7.8	22.8	-9.1	-1.3
		-9.1	143.3						
		-11.9	136.2	219.1	225.9	-6.8	-4.2	-7.7	-0.9
		-19.7	135.5						
		-18.1	127.8						
				219.7	234.1	-14.4	-8[f]	-12	2
		-68.7	88.9						
			121.6						
		-69.6	80.2	217.4	228.6	-11.2	-52[f]	-18	-7
		-26.5	127.8						

Table 1. (Contd.)

B	HB^+	Calcd ΔH_f B	Calcd ΔH_f HB^+	Proton Affinity Calcd	Proton Affinity Exptl[b]	Proton Affinity Error	Exptl[d] $\Delta H_f(B)$	Error in Calcd ΔH_f B	Error in Calcd ΔH_f HB^+
		-24.2	115.9						
				224.8	237.6	-12.8	-13[f]	-14	-1
		42.3	188.5	221.0	223.8	-2.8	46[f]	-4	-1
			188.9	220.6					
			186.5	223.0		-0.8			-3

[a] Ref. 5. [b] Unless otherwise noted, experimental PAs are from Ref. 36a. [c] Ref. 37. While reported in Ref. 36b, these values are not incorporated into the evaluated scale of PAs of Ref. 36. [d] Unless otherwise noted, experimental heats of formation are from Ref. 38. [e] Ref. 15. [f] Estimated in Ref. 36a. [g] Ref. 39. [h] Ref. 36a and references therein.

(1) Carbon Bases. The average unsigned error in the PAs calculated for the carbon bases is 2.9 kcal/mol[16], while that in the heats of formation of the conjugate acids is 3.9 kcal/mol[16]. These are similar to the corresponding error (3.1 kcal/mol) in the heats of formation calculated for the parent hydrocarbons.

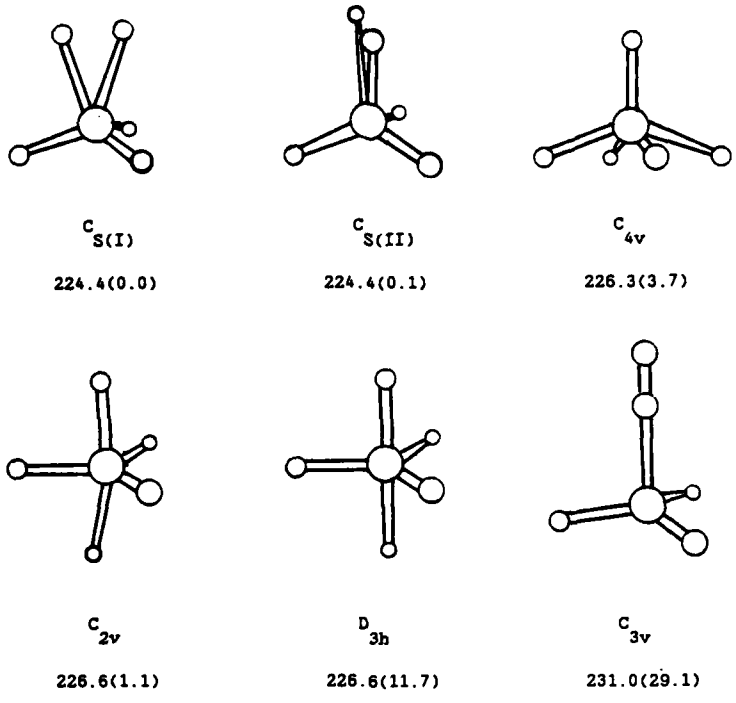
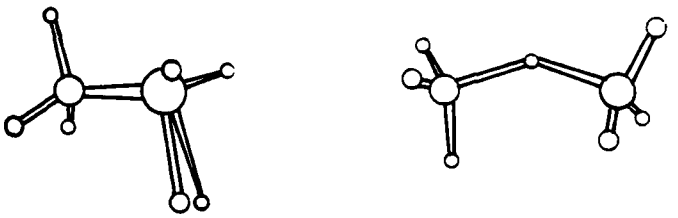


Figure 1. Six structures were optimized for CH_5^+ using AM1. AM1 heats of formation, in kcal/mol, are given below each structure. Values in parentheses are ab initio relative energies from Ref. 8p, also in kcal/mol.

Heats of formation were calculated for six geometries of CH_5^+ cation (Figure 1). In agreement with high level *ab initio* calculations[8p,8nnn,17,18], the $C_{s(I)}$ structure is predicted to be most stable. AM1 calculates the $C_{s(II)}$ structure to be essentially equivalent in stability, its heat of formation differing from that for $C_{s(I)}$ by only 0.001 kcal/mol. While the AM1 results agree with the *ab initio* calculations as to the C_{3v} structure being the least stable isomer and the D_{3h} structure being the second least stable



C_s C_1

208.4(6.8) 221.0(0.0)

Figure 2. Two structures were optimized for $C_2H_7^+$ using AM1. AM1 heats of formation, in kcal/mol, are given below each structure. Values in parentheses are ab initio relative energies from Ref. 17, also in kcal/mol.

isomer, the magnitudes of the differences in heats of formation between these isomers and $C_{s(I)}$ are significantly less for AM1 than for the *ab initio* calculations. Additionally, the C_{2v} structure calculated by AM1 is a very slightly distorted trigonal bipyramid, its heat of formation being essentially the same as that of the D_{3h} structure. The calculated force constants indicate that only the $C_{s(I)}$ and C_{4v} structures correspond to minima on the potential energy (PE) surface, each of the others having at least one negative eigenvalue. Raghavachari *et al* found only the $C_{s(I)}$ structure to be a true minimum[17,19].

Two structures were calculated for $C_2H_7^+$ (Figure 2). In contrast to *ab initio* results[8hh,17] and conclusions based on pulsed high pressure mass spectrometry[20], AM1 predicts the C_s structure alone to correspond to a minimum. The C_1 structure had one negative force constant. Similar results were obtained using MINDO/3[8hh,21].

The calculated heats of formation for protonated alkenes in Table 1 are for classical structures. AM1 fails to predict a nonclassical structure as a true minimum for the ethyl cation[5]. Protonation sites for asymmetric alkenes and alkynes are predicted correctly, as are the preferred para- and ortho-protonation of toluene.

(2) Nitrogen Bases. The average unsigned errors in the PAs calculated for neutral nitrogen bases, and for the heats of formation calculated for them and for their conjugate acids, are 5.8 kcal/mol, 3.9 kcal/mol and 5.3 kcal/mol, respectively. The errors in the calculated PAs and in the heats of formation calculated for the conjugate acids are again similar to that for the neutral bases.

The conjugate acids for all amines except aniline are calculated to be unstable relative to the corresponding bases; i.e. $\delta\Delta H_f(HB^+) - \delta\Delta H_f(B) > 0$. Furthermore, these differences, and the errors in the PAs, generally increase in the sequence $1^o < 2^o < 3^o$. The former trend is probably due to AM1 being parameterized for ground state neutral molecules, in each of which nitrogen has a localized lone pair of electrons. The latter trend can be attributed to the fact that ammonium ions have four atoms bonded to nitrogen, whereas the molecules included in the parameterization basis set had a maximum of three. Moreover, few of the latter had more than one alkyl group bonded to nitrogen.

As a result of these factors, the PAs of amines are underestimated by AM1 in an irregular manner, much as they were by MNDO[11]. The errors are less with AM1 than with MNDO, however, and some relative basicities predicted incorrectly by MNDO are now predicted correctly. For example, AM1 now correctly predicts trimethylamine to be more basic than methylamine, though dimethylamine is still predicted to be less basic than methylamine.

AM1 predicts preferential N-protonation of aniline, in agreement with experiment[22] and theory[23]. AM1 results disagree with the results of Del Bene's high level *ab initio* calculations[8e], however, in predicting cis-ethanimine to be more stable than the trans- isomer.

AM1 correctly predicts quinuclidine to be more basic than its unsaturated analog. The difference between the two calculated PAs is significantly less than the experimental results, however, indicating that AM1 underestimates the inductive effect of the sp^2-carbon in the unsaturated compound.

(3) Oxygen Bases. The average unsigned error in the calculated PAs of the oxygen bases is 6.8 kcal/mol[16], while that in the calculated heats of formation of the cations is 5.4 kcal/mol[16], the corresponding error for the neutral bases being 5.1 kcal/mol.

Experimental trends within each group of compounds (i.e. alcohols, ethers and carbonyl compounds) are accurately reproduced by AM1. As a whole, however, alcohols and ethers are predicted to be weaker bases than experiment indicates by approximately 10 kcal/mol. As in the case of ethanimine, AM1 again differs from Del Bene's *ab initio* results[8g] in predicting preferential *cis*-protonation of acetaldehyde and formamide.

(4) Bifunctional Bases. The unsigned average error in PAs for the six bifunctional bases listed is 9.3 kcal/mol, compared with 10.8 and 2.2 kcal/mol, respectively, for the heats of formation of the bases and conjugate acids. Thus, for these compounds, the heats of formation of the conjugate acids of the bifunctional bases are calculated more accurately than those of the neutral bases. Note that the errors for the neutral bases are all negative and that the errors for three of four n-alkyl bifunctional bases are exceptionally large (see below).

As mentioned previously, these bifunctional bases are of particular interest in that the PAs for five of the six compounds are much larger than those of alkyl amines of comparable polarizability, the exception being hydrazine. The high PAs can be attributed to intramolecular hydrogen bonding in the conjugate acids[1a,24]. For hydrogen bonding to play a role in the protonation of hydrazine, the conjugate acid would have to have a structure analogous to the nonclassical structure of the ethyl cation. No

such structure was found to be a minimum on the AM1 PE surface.

Two geometries were calculated for each of the four n-alkyl bifunctional bases listed in Table 1, one a "cyclic" structure (chain dihedral angles ~0^0) and the other an "extended" structure (chain dihedral angles ~180^0). It was assumed that hydrogen bonding would play a role only in the cyclic structure. Heats of formation are given for both optimized geometries. PAs are calculated using the heats of formation of the more stable conformers.

In the more stable conformer of 1,2-diaminoethane, one amino group is gauche to the other. However, the relative orientation of the two amino groups in the optimized structure precludes hydrogen bonding between them. The lower heat of formation calculated for this conformer relative to that for the conformer in which the two amino groups are anti to each other is a manifestation of the gauche effect[25,26].

If no hydrogen bonding is involved in the protonated base, the difference between the heats of formation calculated for the cyclic and extended conformers of the acid should be approximately the same as for the neutral base, i.e. ΔH_f of the cyclic structure of $H_2NCH_2CH_2NH_3^+$ should be 140.5 kcal/mol. Consequently, AM1 calculates hydrogen bonding to stabilize the cyclic structure of the conjugate acid by 4.3 kcal/mol.

Table 2. Comparison with Experiment of AM1 Hydrogen Bonding Stabilization in the Conjugate Acids of n-Alkyl Bifunctional Bases with Experimental and Ab Initio Results (kcal/mol).

Conjugate Acid	AM1	Exptl	Ab Initio[e]
$H_2N(CH_2)_2NH_3^+$	4.3	6.7[a],9.1[b],12.6[c]	24.8[f],7.7[g]
$H_2N(CH_2)_3NH_3^+$	9.3	14.2[a],14.2[b],20.5[c]	
$HO(CH_2)_3NH_3^+$	7.7[d]	8.8[a],15.2[b]	
$H_2N(CH_2)_4NH_3^+$	14.2	17.9[a],19.8[b]	

[a]Ref. 28. [b]Ref. 29. [c]Ref. 30. [d]Hydrogen bonding stabilization in this acid is underestimated due to hydrogen bonding interaction in the neutral base. See text. [e]Ref. 8jjj. [f]STO-3G//STO-3G. [g]4-31G//STO-3G.

The stabilization due to hydrogen bonding in the conjugate acids of 1,3-diaminopropane, 3-aminopropan-1-ol and 1,4-diaminobutane was calculated in the same manner as for 1,2-diaminoethane. The AM1 results are given in Table 2, along with experimental estimates by Meot-Ner *et al*[28], Buschek *et al*[29], and Yamdagni *et al*[30], and the *ab initio* results of Houriet *et al*[8jjj]. The AM1 values are similar to the most recent experimental values, i.e. those of Meot-Ner[28,31], the N-H···N hydrogen bond being somewhat underestimated by AM1. The calculated value for 3-aminopropan-1-ol does not indicate the full extent of hydrogen bonding in the acid since hydrogen bonding is also involved in the neutral base[29,32]. The preferred N-protonation as opposed to O-protonation is in agreement with previous theoretical calculations and experimental results[8jjj].

Table 3. Comparison of Errors in AM1 Heats of Formation for n-Alkanes, n-Alkyl Amines and n-Alkyl Alcohols (kcal/mol).

	$CH_3(CH_2)_nCH_3$			$CH_3(CH_2)_nNH_2$			$CH_3(CH_2)_nOH$		
	AM1 ΔH_f	Exptl ΔH_f[b]	Error	AM1 ΔH_f	Exptl ΔH_f[b]	Error	AM1 ΔH_f	Exptl ΔH_f[b]	Error
n=0	-17.4[a]	-20.0	2.6	-7.4[a]	-5.5	-1.9	-57.0[a]	-48.2	-8.8
n=1	-24.3[a]	-25.0	0.7	-15.1[a]	-11.3	-3.8	-62.7[a]	-56.2	-6.5
n=2	-31.1[a]	-30.0	-1.1	-22.1[a]	-16.8	-5.3	-70.6[a]	-61.0	-9.6
n=3	-37.9[a]	-35.1	-2.8	-28.9	-22.0	-6.9	-77.3	-65.7	-11.6
n=4	-44.1	-39.9	-4.2	-35.1	-26[c]	-9	-84.1	-70.4	-13.7

[a] Ref. 5. [b] Ref. 38. [c] Estimated in Ref. 36a.

Despite the good predictions of the strengths of the hydrogen bonds, the errors in the calculated PAs for the n-alkyl bifunctional bases are exceptionally large. The errors in the corresponding PAs are, moreover, due largely to errors in the calculated heats of formation of the neutral bases rather than in those of the conjugate acids. The reason for this is a combination of systematic errors, one reflecting increasing alkyl chain lengths and the other substitution of nitrogen or oxygen for a primary carbon atom. Table 3 lists the errors in AM1 heats of formation for a series of n-alkanes, n-alkyl amines and n-alkyl alcohols. The errors for the alkanes change by an average of -1.7 kcal/mol for each additional methylene group in the alkyl chain[33]. Likewise substitution of a nitrogen atom for a primary carbon changes the

error for an n-alkyl amine by an average of -4.4 kcal/mol relative to the error for the corresponding alkane. The corresponding change in the error due to replacement of a primary carbon by oxygen is -9.1 kcal/mol. These errors are approximately additive, and the large errors in the heats of formation of the bifunctional bases can be explained in terms of them. The corresponding errors for substitution for a primary carbon bonded to a tertiary carbon are less, being -1.8 and -5.2 kcal/mol for amino and hydroxyl substitution, respectively[34]. No such correlation exists for substitution for a secondary or tertiary carbon, or for substitution for a primary carbon bonded to a quaternary carbon.

Catalán concluded in a recent INDO study[35] that 1,8-diaminonaphthalene is a nitrogen base, the PA calculated as 231.1 kcal/mol for amine protonation as opposed to 216.2 kcal/mol for ring protonation. The actual magnitude of the preference for protonation on nitrogen remained uncertain, due to the well known tendency of INDO to overestimate the strengths of hydrogen bonds[35]. AM1, however, predicts a higher proton affinity for ring protonation. Neither the AM1 results for naphthalene, nor those for aniline, provide any indication that AM1 overestimates the stabilities of the ring protonated cations or underestimates those of the amine protonated ones. However, AM1 underestimates the strength of the N-H···N intramolecular hydrogen bond by at least 2.4 - 4.9 kcal/mol, as indicated by the calculations for the n-alkyl diamines. Consequently, the proton affinity for amine protonation

Table 4. Comparison of Ab Initio and AM1 Proton Affinities[9] (kcal/mol).

	6-31G*//3-21G		AM1
Molecule	PA[a]	Error[b]	Error[c]
NH_3	218.4	14.4	5.3
CH_3NH_2	228.5	14.4	-3.0
$(CH_3)_2NH$	235.2	14.6	-9.7
HCN	178.1	6.7	12.0
CH_3CN	194.0	5.6	2.0
H_2O	173.3	6.9[d]	-2.0[d]
CH_3OH	189.5	7.6	-10.0
$(CH_3)_2O$	198.6	6.5	-14.7
H_2CO	181.4	9.7	2.7
CH_3CHO	194.0	7.4	-1.7
Avg. Unsigned Error		9.4	6.3

[a] Ref. 8jj. [b] Errors are calculated using experimental values from Table 1. [c] From Table 1. [d] Based on the average of the two experimental values given in Table 1.

is probably 0.4 - 2.9 kcal/mol higher than that for ring protonation. The small difference between N-protonation and ring protonation seems quite reasonable, given that aniline undergoes N-protonation[22] in the gas phase while *m*-diaminobenzene[22a,c] and 1-aminonaphthalene[24] protonate in the ring.

(5) *Ab Initio* Calculations. The results of high level *ab initio* calculations[8jj] of PAs for ten of the molecules listed in Table 1 are summarized in Table 4. While the errors naturally fluctuate, the AM1 values are, on average, more accurate by a significant margin.

B. Deprotonation Enthalpies

The deprotonation enthalpy (DPE) of compound HB is the heat of reaction for loss of a proton to form the conjugate base;

$$HB \rightarrow H^+ + B^-$$

$$DPE(HB) = \Delta H_f(H^+) + \Delta H_f(B^-) - \Delta H_f(HB) \qquad (3)$$

The DPE of a compound is thus equal to the PA of its conjugate base. Which terminology is chosen (PA or DPE) depends on the charges present, PA and DPE being regarded primarily as properties of neutral molecules.

Calculated DPEs for 80 compounds are given in Table 5, along with corresponding experimental values. Also included are the errors in the calculated heats of formation of the conjugate bases, $\delta\Delta H_f(B^-)$, calculated using the equation;

$$\delta\Delta H_f(B^-) = \delta DPE(HB) + \delta\Delta H_f(HB) \qquad (4)$$

where δDPE(HB) and $\delta\Delta H_f$(HB) are the errors in the calculated DPE and heat of formation of HB, respectively. The listed compounds include 46 carbon acids (the lowest DPE is for a hydrogen bonded to a carbon), 10 nitrogen acids and 24 oxygen acids.

(1) Carbon Acids. The average unsigned error in calculated DPEs for the carbon acids is 8.0 kcal/mol[16]. The values for CH_4, C_2H_4, HCN, and the alkynes are exceptionally poor as a result of large errors in the calculated heats of formation of the corresponding anions. When these eight compounds are excluded from the statistical analysis, the average unsigned error is reduced to 5.7 kcal/mol[16]. The average unsigned error in the heats of formation of the 38 anions is also 5.7 kcal/mol[16]. These errors compare favorably with the error of 5.1 kcal/mol for the 38 corresponding neutral acids.

The poor results for the CH_3^- and CN^- anions can be attributed to the failure of AM1 to allow for orbital expansion on atoms bearing large negative charges[5]. Similar problems were encountered in calculations for anions using MNDO[11]. The calculated heats of formation are expected to be, and are, too positive whenever the formal charge in an anion is largely concentrated on a single atom. This is also the case for HO^-[5] and H_2N^- anions (see below). The same problem arises, as expected, in the case of allenyl anion[40] and propynyl anion. Both are incorrectly[41] predicted to be less stable than propargyl anion ($^-CH_2C{\equiv}CH$, C_{2v}).

While the negative charge in the allenyl or propynyl anion is concentrated largely on a single carbon atom, in propargyl it is dispersed over an allylic system.

Experiment[43] indicates that the propynyl anion is lower in energy than allenyl anion by 3 kcal/mol while *ab initio* calculations[8w,44,45] imply the allenyl anion is lower than propargyl by 7-9 kcal/mol. Thus the difference between the propargyl and propynyl anions is ca 11 kcal/mol. Using the experimental[42] value for the heat of formation for the latter, we arrive at an estimate (68 kcal/mol) for that of propargyl anion, in good agreement with AM1 (64.5 kcal/mol). This supports the interpretation given above of the errors for allenyl and propynyl.

A similar situation is expected in the case of propene and its derivatives. While allyl anion is correctly predicted to be the most stable isomer, the difference between it and 1- or 2-propenyl anion is probably overestimated by AM1. The relative order of stability predicted by AM1 agrees with that given by MNDO[11] and by the *ab initio* calculations of Boerth and Streitwieser[8mm], except in that the E- and Z- isomers of 1-propenyl anion are predicted by AM1 to have essentially identical energies. AM1 also reproduces the effects of substituents at the central carbon atom, attributed by Bartmess and Burnham to be primarily polar in nature[46].

Table 5. Comparison with Experiment of AM1 Heats of Formation and Deprotonation Enthalpies (kcal/mol).

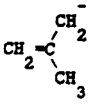
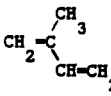
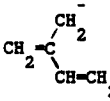
HB	B^-	Calcd ΔH_f HB	Calcd ΔH_f B^-	Deprotonation Enthalpy Calcd	Deprotonation Enthalpy Exptl[b]	Deprotonation Enthalpy Error	Exptl ΔH_f(HB)[c]	Error in Calcd ΔH_f HB	Error in Calcd ΔH_f B^-
Carbon Acids									
CH_4	CH_3^-	-8.8[a]	57.7	433.7	416.6	17.1	-17.8	9.0	26.1
CH_3CH_3	$CH_3CH_2^-$	-17.4[a]	34.5	419.1	421.0	-1.9	-20.0	2.6	0.7
$CH_3CH_2CH_3$	$CH_3\overset{-}{C}HCH_3$	-24.3[a]	16.7	408.2	419.0	-10.8	-25.0	0.7	-10.1
$(CH_3)_3CH$	$(CH_3)_3C^-$	-29.4[a]	3.1	399.7	414.0	-14.3	-32.1	2.7	-11.6
		17.8[a]	67.6	417.0	412.0	5.0	12.7	5.1	10.1
		8.6[a]	27.6	388.2	390.7	-2.5	4.8	1.8	-0.7
			50.6	411.2					
			58.4	417.0					
			56.7	417.3					
$CH_2=C(CH_3)_2$		-1.2[a]	20.7	389.1	390.3	-1.2	-4.0	2.8	1.6
		23.3	42.5	386.4	385.0	1.4	18.0	5.3	6.7
		37.1[a]	25.2	355.3	356.1	-0.8	32.1	5.0	4.2

Table 5. (Contd.)

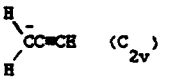
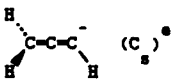
HB	B^-	Calcd ΔH_f HB	Calcd ΔH_f B^-	Deprotonation Enthalpy Calcd	Deprotonation Enthalpy Exptl[b]	Deprotonation Enthalpy Error	Exptl ΔH_f(HB)[c]	Error in Calcd ΔH_f HB	Error in Calcd ΔH_f B^-
		38.3[a]	39.6	368.5	373.9	-5.4	43.2	-4.9	-10.3
$CH_2{=}CH_2$		16.5[a]	67.8	418.5	406.0	12.5	12.5	4.0	16.5
HC≡CH	$HC{\equiv}C^-$	54.6[a]	89.1	401.5	375.4 384.5	26.1 17.0	54.5	0.3	26.4
$CH_3C{\equiv}CH$	$CH_3C{\equiv}C^-$	43.4[a]	74.9	398.7	379.7	19.0	44.2	-0.8	18.2
	 (C_{2v})		64.5	388.3					
	 (C_s)[e]		73.1	396.9[f]					
$CH_3(CH_2)_2C{\equiv}CH$	$CH_3(CH_2)_2C{\equiv}C^-$	30.6	60.8	397.4	378.4	19.0	34.7[g]	-4.1	14.9
$(CH_3)_3CC{\equiv}CH$	$(CH_3)_3CC{\equiv}C^-$	30.8	60.3	396.7	376.7	20.0	25.1[g]	5.7	25.7
$C_6H_5C{\equiv}CH$	$C_6H_5C{\equiv}C^-$	76.5	94.3	385.0	370.4	14.6	78.3[g]	-1.8	12.8
C_6H_6	$C_6H_5^-$	22.0[a]	57.4	402.6	398.8	3.8	19.8	2.2	6.0
$C_6H_5CH_3$	$C_6H_5CH_2^-$	14.5[a]	20.8	373.5	379.4	-5.9	12.0	2.5	-3.4
$4\text{-}CH_3C_6H_4CH_3$	$4\text{-}CH_3C_6H_4CH_2^-$	6.8	12.2	372.6	380.5	-7.9	4.3	2.5	-5.4
$C_6H_5CH_2CH_3$	$C_6H_5\bar{C}HCH_3$	8.8[a]	10.2	368.6	378.3	-9.7	7.2	1.6	-8.1
$C_6H_5CH(CH_3)_2$	$C_6H_5\bar{C}(CH_3)_2$	4.7	2.5	365.0	377.5	-12.5	1.0	3.7	-8.8
$(C_6H_5)_2CH_2$	$(C_6H_5)_2CH^-$	42.1	29.2	354.3	364.5	-10.2	33.2	8.9	-1.3

Table 5. (Contd.)

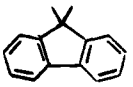
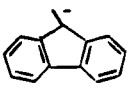
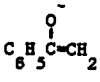
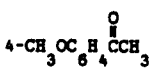
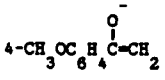
HB	B^-	Calcd ΔH_f HB	Calcd ΔH_f B^-	Deprotonation Enthalpy Calcd	Deprotonation Enthalpy Exptl[b]	Deprotonation Enthalpy Error	Exptl ΔH_f(HB)[c]	Error in Calcd ΔH_f HB	Error in Calcd ΔH_f B^-
		54.4	36.1	348.9	353.3	-4.4	45[h]	9	5
HCN	CN^-	31.0[a]	44.0	380.2	353.1	27.1	32.3[i]	-1.3	25.8
CH_3CN	CH_2CN^-	19.3[a]	30.8[a]	378.7	372.1	6.6	15.4	3.9	10.5
CH_3CH_2CN	CH_3CHCN^-	13.0[a]	17.9	372.1	373.7	-1.6	12.3	0.7	0.9
$(CH_3)_2CHCN$	$(CH_3)_2CCN^-$	8.5	7.5	366.2	373.8	-7.6	5.6	2.9	-4.7
		48.8	56.1	374.5	374.1	0.4	43.5	5.3	5.7
CH_3OCH_2CN	CH_3OCHCN^-	-17.1	-16.0	368.3	371.4	-3.1	-13.0[g]	-4.1	-7.2
$C_6H_5CH_2CN$	$C_6H_5CHCN^-$	46.7	26.9	347.4	353.3	-5.9	44.5[g]	2.2	-3.7
		38.0	43.8	373.0	371.6	1.4	36.0[g]	2.0	3.4
CH_3OCH_3	$CH_3OCH_2^-$	-53.2[a]	-8.6	411.8	407.0	4.8	-44.0	-9.2	-4.4
CH_3CHO	CH_2CHO^-	-41.6[a]	-37.2	371.6	366.4	5.2	-39.7	-1.9	3.3
	CH_3CO^-		-14.1	394.7					
CH_3CH_2CHO	CH_3CHCHO^-	-48.0[a]	-49.9	365.3	365.7	-0.4	-44.4	-3.6	-4.0
		-49.2[a]	-43.6	372.8	368.8	4.0	-51.9	2.7	6.7
		-15.0	-15.9	366.3	363.2	3.1	-20.7	5.7	8.8
		-52.7	-55.3	364.6	362.8	1.8	-59.8	7.1	8.9

Table 5. (Contd.)

HB	B^-	Calcd ΔH_f HB	Calcd ΔH_f B^-	Deprotonation Enthalpy Calcd	Deprotonation Enthalpy Exptl[b]	Deprotonation Enthalpy Error	Exptl ΔH_f(HB)[c]	Error in Calcd ΔH_f HB	Error in Calcd ΔH_f B^-
$C_6H_5\overset{O}{\overset{\|}{C}}CH_2CH_3$	$C_6H_5\overset{O^-}{\overset{\|}{C}}{=}CHCH_3$	-20.8	-26.5	361.5	362.4	-0.9	-26.0	5.2	4.3
$C_6H_5CH_2\overset{O}{\overset{\|}{C}}CH_3$	$C_6H_5CH{=}\overset{O^-}{\overset{\|}{C}}CH_3$	-20.3	-41.0	346.5	352.5	-6.0	-24.1	3.8	-2.2
$CH_3\overset{O}{\overset{\|}{C}}CH_3$	$CH_2{=}\overset{O^-}{\overset{\|}{C}}CH_3$	-96.4[a]	-94.2	369.4	371.0	-1.6	-98.4	2.0	0.4
	$CH_3\overset{O}{\overset{\|}{C}}CH_2^-$		-63.3	400.3					
$CH_3\overset{O}{\overset{\|}{C}}CH_2\overset{O}{\overset{\|}{C}}CH_3$	$CH_3\overset{O}{\overset{\|}{C}}CH{=}\overset{O^-}{\overset{\|}{C}}CH_3$	-131.1	-141.8	356.5	350.3[j]	6.2	-137.9[k]	6.8	13.0
	$CH_2{=}\overset{O^-}{\overset{\|}{C}}CH_2\overset{O}{\overset{\|}{C}}CH_3$		-138.7	359.6					
	$CH_3\overset{O}{\overset{\|}{C}}CH_2\overset{O^-}{\overset{\|}{C}}{=}CH_2$		-133.6	364.7					
$CH_3\overset{O}{\overset{\|}{C}}N(CH_3)_2$	$CH_2{=}\overset{O^-}{\overset{\|}{C}}N(CH_3)_2$	-33.6	-33.3	367.5	373.5	-6.0	-55.6[h]	22.0	16.0
CH_3NO_2	$CH_2{=}NO_2^-$	-9.9[a]	-29.2[a]	347.9	358.7 357.6	-10.8 -9.7	-17.8	7.9	-2.9
$CH_3CH_2NO_2$	$CH_3CH{=}NO_2^-$	-16.8	-39.0	345.0	358.1	-13.1	-24.5	7.7	-5.4
$(CH_3)_2CHNO_2$	$(CH_3)_2C{=}NO_2^-$	-21.5	-46.0	342.7	358.2	-15.5	-33.2	11.7	-3.8
$(CH_3)_3CCH_2NO_2$	$(CH_3)_3CCH{=}NO_2^-$	-30.7	-53.0	344.9	357.3	-12.4	-45.2[g]	14.5	2.1
Nitrogen Acids									
NH_3	H_2N^-	-7.3[a]	52.5	427.0	403.6	23.4	-11.0[m]	3.7	27.1
CH_3NH_2	CH_3NH^-	-7.4[a]	33.1[a]	407.7	403.2	4.5	-5.5	-1.9	2.6

Table 5. (Contd.)

HB	B^-	Calcd ΔH_f HB	Calcd ΔH_f B^-	Deprotonation Enthalpy Calcd	Deprotonation Enthalpy Exptl[b]	Deprotonation Enthalpy Error	Exptl ΔH_f(HB)[c]	Error in Calcd ΔH_f HB	Error in Calcd ΔH_f B^-
	$^-CH_2NH_2$		40.3	414.9					
$CH_3CH_2NH_2$	$CH_3CH_2NH^-$	-15.1[a]	27.4	409.7	399.4	10.3	-11.3	-3.8	6.5
$C_6H_5NH_2$	$C_6H_5NH^-$	20.7[a]	19.4	365.9	367.1	-1.2	20.8	-0.1	-1.3
$3\text{-}CH_3C_6H_4NH_2$	$3\text{-}CH_3C_6H_4NH^-$	13.0	12.1	366.3	367.5	-1.2	12.9[g]	0.1	-1.1
$4\text{-}CH_3C_6H_4NH_2$	$4\text{-}CH_3C_6H_4NH^-$	12.9	10.9	365.2	368.2	-3.0	12.9[g]	0.0	-3.0
$4\text{-}CH_3OC_6H_4NH_2$	$4\text{-}CH_3OC_6H_4NH^-$	-16.6	-21.2	362.6	368.0	-5.4	-16.5[g]	-0.1	-5.5
$(CH_3)_2NH$	$(CH_3)_2N^-$	-5.6[a]	22.4[a]	395.2	396.4	-1.2	-4.4	-1.2	-2.4
		39.9[a]	28.1[a]	355.4	360.8	-5.4	25.9	14.0	8.6
$C_6H_5NHCCH_3$ (O)	$C_6H_5NCCH_3$ (O)	-15.0	-34.6	347.6	352.7	-5.1	-30.8	15.8	10.7
Oxygen Acids									
H_2O	HO^-	-59.2[a]	-14.1[a]	412.3	390.8	21.5	-57.8[m]	-1.4	20.1
CH_3OH	CH_3O^-	-57.0[a]	-38.8[a]	385.4	379.2	6.2	-48.2	-8.8	-2.6
					381.4[d]	4.0			-4.8
CH_3CH_2OH	$CH_3CH_2O^-$	-62.7[a]	-45.8[a]	384.1	376.1	8.0	-56.2	-6.5	1.5
					378.3[d]	5.8			-0.7
$CH_3(CH_2)_2OH$	$CH_3(CH_2)_2O^-$	-70.6[a]	-53.7	384.1	374.7	9.4	-61.0	-9.6	-0.2
					376.9[d]	7.2			-2.4
$(CH_3)_2CHOH$	$(CH_3)_2CHO^-$	-67.7[a]	-50.7	384.2	374.1	10.1	-65.2	-2.5	7.6
					376.3[d]	7.9			5.4
$(CH_3)_2CHCH_2OH$	$(CH_3)_2CHCH_2O^-$	-74.0	-59.5	381.7	373.4	8.3	-67.9	-6.1	2.2
					375.6[d]	6.1			0.0

Table 5. (Contd.)

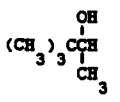
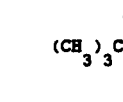
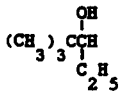
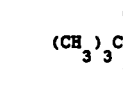
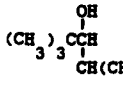
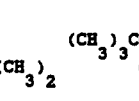
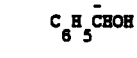
HB	B^-	Calcd ΔH_f HB	Calcd ΔH_f B^-	Deprotonation Enthalpy Calcd	Deprotonation Enthalpy Exptl[b]	Deprotonation Enthalpy Error	Exptl ΔH_f(HB)[c]	Error in Calcd ΔH_f HB	Error in Calcd ΔH_f B^-
$(CH_3)_3COH$	$(CH_3)_3CO^-$	-71.6[a]	-54.2	384.6	373.3	11.3	-74.7	3.1	14.4
					375.5[d]	9.1			12.2
$(CH_3)_3CCH_2OH$	$(CH_3)_3CCH_2O^-$	-76.5	-63.5	380.2	371.8	8.4	-76.1[g]	-0.4	8.0
					374.0[d]	6.2			5.8
		-79.7	-66.1	380.8	370.7	10.1	-83.8[g]	4.1	14.2
		-86.1	-73.2	380.1	369.6	10.5	-88.7[g]	2.6	13.1
		-88.6	-78.4	379.4	368.5	10.9	-93.6[g]	5.0	15.9
$((CH_3)_3C)_2CHOH$	$((CH_3)_3C)_2CHO^-$	-86.3	-75.3	378.2	367.3	10.9	-99.1[g]	12.8	23.7
$CH_3O(CH_2)_2OH$	$CH_3O(CH_2)_2O^-$	-103.5	-92.2	378.5	372.5	6.0	-87.4[g]	-16.1	-10.1
$C_6H_5CH_2OH$	$C_6H_5CH_2O^-$	-31.2	-21.7	376.7	369.6	7.1	-24.0	-7.2	-0.1
			-29.5	368.9					
C_6H_5OH	$C_6H_5O^-$	-22.2[a]	-41.0[a]	348.4	349.8	-1.4	-23.0	0.8	-0.6
$4\text{-}CH_3C_6H_4OH$	$4\text{-}CH_3C_6H_4O^-$	-29.8	-49.2	347.8	350.9	-3.1	-30.0	0.2	-2.9
HCO_2H	HCO_2^-	-97.4[a]	-110.0[a]	354.6	345.2	9.4	-90.5	-6.9	2.5
CH_3CO_2H	$CH_3CO_2^-$	-103.0[a]	-116.0[a]	354.2	348.5	5.7	-103.4	0.4	6.1
$CH_3CH_2CO_2H$	$CH_3CH_2CO_2^-$	-108.0[a]	-122.3	352.9	347.3	5.6	-108.4	0.4	6.0
$C_6H_5CO_2H$	$C_6H_5CO_2^-$	-68.0[a]	-86.9	348.3	338.3	10.0	-70.3	2.3	12.3
$CH_3CH{=}NOH$	$CH_3CH{=}NO^-$	-8.4	-22.8	352.8	366.2	-13.4	-4.7[l]	-3.7	-17.1

Table 5. (Contd.)

HB	B^-	Calcd ΔH_f HB	Calcd ΔH_f B^-	Deprotonation Enthalpy Calcd	Deprotonation Enthalpy Exptl[b]	Deprotonation Enthalpy Error	Exptl ΔH_f(HB)[c]	Error in Calcd ΔH_f HB	Error in Calcd ΔH_f B^-
$(CH_3)_2C{=}NOH$	$(CH_3)_2C{=}NO^-$	-15.3	-31.3	351.2	366.7	-15.5	-13.0[g]	-2.3	-17.8
$(CH_3)_3CCH{=}NOH$	$(CH_3)_3CCH{=}NO^-$	-21.9	-37.4	351.7	364.6	-12.9	-23.2[g]	1.3	-11.6
$C_6H_5CH{=}NOH$	$C_6H_5CH{=}NO^-$	24.4	-5.5	337.3	355.1	-17.8	25.8[g]	-1.4	-19.2

[a] Ref. 5. [b] Unless otherwise noted, experimental DPEs are from Ref. 52. [c] Unless otherwise noted, experimental heats of formation are from Ref. 38. [d] Ref. 54. [e] Ref. 40. [f] Calculated for the formation of propyne upon protonation. [g] Estimated in Ref. 52. [h] Ref. 36a and references therein. [i] Ref. 15 [j] Ref. 1b. See text. [k] Estimated by group additivity method, Ref. 53. Value for $C-(H)_2(CO)(O)$ approximated as that for $C-(H)_2(O)(C)$. [l] Ref. 52 and references therein. [m] Ref. 39.

The errors in the DPEs calculated for the nitro alkanes are also relatively large. These, however, are due primarily to errors in the heats of formation calculated for the neutral acids[5] rather than in those for their conjugate bases.

In all but one case AM1 overestimates the stabilizing effect of a methyl or phenyl substituent at the anionic center in a carbanion, the exception being the 2-nitro-2-propyl anion. The effect is most significant when the parent anion is primary, the error being approximately -8 kcal/mol for methyl substitution and -10 kcal/mol for phenyl substitution[47].

The experimental evidence[48] and *ab initio*[7] calculations agree with the AM1 prediction, that the α-hydrogens of carbonyl compounds are more acidic than the acyl hydrogen. The difference between the corresponding DPEs in acetaldehyde is calculated by AM1 to be 23 kcal/mol, in reasonable agreement with the value (29 kcal/mol) derived from 4-31 + G//4-31 + G and MP2/4-31 + G//4-31 + G calculations[7]. Additionally, AM1 predicts the lowest DPE in acetonyl acetate to be that for the methene hydrogens, in agreement with an earlier assignment[1b]. However, due to uncertainty concerning the change in entropy accompanying the loss of a methene hydrogen, neither the deprotonation site nor the experimental value of the DPE is firmly established[49].

(2) Nitrogen Acids. The average unsigned error in the DPEs calculated for the nitrogen acids listed in Table 5 is 6.1 kcal/mol. If NH_3 is eliminated for the reasons indicated above, the error is reduced to 4.1 kcal/mol, and that for the heats of formation of the anions to 4.6 kcal/mol. These values again compare favorably with the corresponding error (4.1 kcal/mol) in the heats of formation of the neutral acids.

The relative acidities of methylamine and ethylamine are predicted incorrectly by AM1. This error is due to two systematic errors discussed above in detail, one being the error accompanying methyl substitution at an anionic site while the other involves increasing alkyl chain length. The prediction that methylamine

behaves as a nitrogen acid rather than a carbon acid coincides with chemical intuition and with the *ab initio* results of Lohr and Ponas[8d].

The acidities of *p*-methyl- and *p*-methoxyaniline are predicted incorrectly relative to aniline. These errors can again be attributed to the overestimation of the stabilizing effects of methyl groups attached to negatively charged carbon atoms.

(3) Oxygen Acids. For the oxygen acids listed in Table 1, the average unsigned error in the DPEs is 9.4 kcal/mol[16]. The error in the case of H_2O has already been discussed. The errors for oximes are also significantly larger than those for the other compounds (see below). Excluding water and the oximes from the statistical analysis, the average unsigned error for the remaining 19 DPEs falls to 7.6 kcal/mol[16], and that in the heats of formation of the conjugate bases to 7.4 kcal/mol[16]. These are larger than the corresponding errors for the carbon acids or nitrogen acids, and they are also larger than the average unsigned error (5.0 kcal/mol) in the heats of formation of the parent acids. Since the errors associated with compounds containing oxygen are, in general, larger than those for hydrocarbons or nitrogen-containing compounds[5], and since the charge on oxygen in an oxygen-containing anion is usually large, it is not surprising that the AM1 errors for such species are larger than usual.

The large errors in the DPEs of oximes are due to the overestimation of the stability of the conjugate bases. Given the tendency of AM1 to overestimate the stabilization associated with phenyl and methyl substitution on the anionic center (see above), the pattern of errors in the heats of formation of the variously substituted oxime conjugate bases is consistent with significant charge buildup on the methene carbon (i.e. errors in heats of formation are large and negative with methyl and phenyl substitution on the methene carbon, but decrease significantly with t-butyl substitution). Analysis of atomic charges in ethanal oxime indicates that 0.40 of the formal charge of the anion is on the methene carbon, with the methyl group absorbing another 0.13. Additionally, the coefficient of the methene carbon p-orbital in the HOMO increases from 0.57 in the neutral acid to 0.75 in the anion. These results point to an overestimation of charge delocalization by AM1 through a conjugated system terminated by a methyl or phenyl group, similar to the effect noted with the anilines.

While the decrease in the DPEs of aliphatic alcohols with increasing size of the alkyl group is reproduced qualitatively by AM1, the calculated difference (7.2 kcal/mol) between the highest DPE (CH_3OH) and the lowest DPE (t-Bu_2CHOH) is less than that observed (11.9-14.1 kcal/mol). Moreover, the order of decreasing DPEs is not in agreement with experiment. While the scatter in the errors of the calculated heats of formation of the neutral alcohols may be partly responsible, the major problem seems to lie in the

errors increasing with the size of the alkyl groups. Thus AM1 does not fully reproduce the charge-induced dipole stabilization of the anions to which the trend in gas-phase acidities in alcohols has been attributed[50,51].

The benzylic hydrogens are calculated to be more acidic than the hydroxyl hydrogen in benzyl alcohol. Applying a 10 kcal/mol correction to the heat of formation of the carbon anion to compensate for the overestimation by AM1 of the added stability due to phenyl substitution at an anionic center, the hydroxyl hydrogen becomes the more acidic, as is observed[42].

(4) Comparison with *Ab Initio Results*. Table 6 lists *ab initio* DPEs, without (4-31 + G//4-31 + G) and with (MP2/4-31 + G//4-31 + G) allowance for electron correlation by second-order Moller-Plesset (MP) perturbation theory[7], for nine of the compounds included in Table 1. Five other molecules included in both studies (CH_4, NH_3, H_2O, HCN, HCCH) have not been included in Table 6 since it has already been recognized that AM1 cannot describe the corresponding conjugate bases adequately.

The basis set in the *ab initio* study includes diffuse AOs, known[7] to be essential for the proper description of anions. While a few *ab initio* values are better than the corresponding AM1 ones, the overall average error is less for AM1.

Table 6. Comparison of Ab Initio and AM1 Deprotonation Enthalpies[9] (kcal/mol).

	4-31+G//4-31+G		MP2/4-31+G//4-31+G		AM1
Molecule	DPE[a]	Error[b]	DPE[a]	Error[b]	Error[c]
C_2H_6	439.1	18.1	432.0	11.0	-1.9
C_2H_4	423.8	17.8	417.9	11.9	12.5
CH_2CHCH_3	405.5	14.8	399.1	8.4	-2.5
CH_3CN	386.1	14.0	383.4	11.3	6.6
CH_3CHO	374.5	8.1	369.2	2.8	5.2
CH_3NH_2	422.6	19.4	410.7	7.5	4.5
CH_3OH	393.1	12.8[d]	381.4	1.1[d]	5.1[d]
CH_3CH_2OH	391.7	14.5[d]	378.7	1.5[d]	6.9[d]
HCOOH	346.9	1.7	337.3	-7.9	9.4
Avg. Unsigned Error		13.5		7.0	6.0

[a] Ref. 7. [b] Errors are calculated using experimental values from Table 5. [c] From Table 5. [d] Based on the average of the two experimental values quoted in Table 5.

Conclusions

With a few exceptions, AM1 seems to be an effective method for studying processes involving deprotonation or protonation of neutral molecules. The errors in the calculated DPEs and PAs, as well as in the calculated heats of formation of deprotonated and protonated species, are comparable with those in the calculated heats of formation of the neutral precursors. Intramolecular hydrogen bonding in protonated bifunctional bases is also

effectively reproduced.

The main problems encountered with AM1 involve small anions, in which the charge is largely concentrated on one atom, and anions formed by the deprotonation of oximes. Systematic errors accompany introduction of methyl or phenyl substituents at anionic centers. Systematic errors also arise in the extension of alkyl chains by addition of methene groups and in substitution of amine and hydroxyl groups for methyl groups bonded to secondary or tertiary carbons in neutral molecules. The errors involved in the deprotonation of alcohols and protonation of amines are not, however, totally systematic. As a result, the relative DPEs of alcohols and PAs of amines are not accurately reproduced by AM1.

The comparisons in Tables 4 and 6 suggest that AM1 performs as well here as do quite high level *ab initio* procedures[55]. The fact that the comparisons refer only to a few simple cases is due simply to the dearth of relevant *ab initio* data. Calculations at this level, if carried out with full geometry optimization, become very expensive for larger molecules. Since the accuracy achieved by AM1 is sufficient for the results to be chemically useful and since it can be used to study reactions of quite large molecules at moderate cost, AM1 should prove useful as an aid in interpretating proton transfer in chemistry and biochemistry.

References

1.For a recent review including discussion of experimental aspects and results, see: (a) Aue, D. H.; Bowers, M.T. in *Gas Phase Ion Chemistry*, Bowers, M.T., Ed; Academic Press: New York, 1979; Vol. 2; Chapter 9. (b) Bartmess, J.E.; McIver, R. T., Jr. *Ibid.* Chapter 11.

2. (a) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69. (b) Hall, G. G. *Proc. R. Soc. London, Ser. A* **1951**, *205*, 541.

3. Bingham, R. C.; Dewar, M. J. S.; Lo, D. H. *J. Am. Chem. Soc.* **1975**, *97*, 1285, 1294, 1302, 1307.

4. Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977** *99*, 4899, 4907.

5. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

6 (a) Dewar, M. J. S.; Ford, G. P. *J. Am. Chem. Soc.* **1979**, *101*, 5558. (b) Gregory, A. R.; Paddon-Row, M. N. *Ibid.* **1976**, *98*, 7521. (c) Boyd, D. B. *J. Phys. Chem.* **1978**, *82*, 1407.

7. Chandrasekhar, J.; Andrade, J.G.; Schleyer, P.v.R., *J. Am. Chem. Soc.* **1981**, *103*, 5609, 5612.

8. (a) Ikuta, S. *J. Comput. Chem.* **1984**, *5*, 374. (b) Lee, C.C.; Hass, E.C.; Obafemi, C.A.; Mezey, P.G. *Ibid.* **1984**, *5*, 190. (c) Lohr, L.L.; Schlegel, H.B.; Morokuma, K. *J. Phys. Chem.* **1984**, *88*, **1981**. (d) Lohr, L.L.; Ponas, S.H. *Ibid.* **1984**, *88*, 2992. (e) Del Bene, J.E. *J. Comput. Chem.* **1984**, *5*, 381. (f) Del Bene, J.E.; Mette, D.D.; Frisch, M.J.; Luke, B.T.; Pople, J.A. *J. Phys. Chem.* **1983**, *87*, 3279. (g) Del Bene, J.E. *Chem. Phys. Lett.* **1983**, *94*, 213. (h) Jorgensen, W.L.; Ibraham, M. *J. Comput. Chem.* **1981**, *2*, 7. (i) Hinde, A.; Radom, L. *Ibid.* **1980**, *1*, 118. (j) Heidrich, D.; Volkmann, D.; Zurawski, B. *Chem. Phys. Lett.* **1981**, *80*, 60. (k) Tel, L.M.; Wolfe, S.; Csizmadia, I.G. *J. Chem. Phys.* **1973**, *59*, 4047. (l) Owens, P.H.; Wolf, R.A.; Streitwieser, A., Jr. *Tetrahedron Lett.* **1970**, 3385. (m) Hehre, W.J.; Pople, J.A. *Tetrahedron Lett.* **1970**, 2959. (n) Edgecombe, K.E.; Boyd, R.J. *Can. J. Chem.* **1984**, *62*, 2887. (o) Clark, T.; Chandrasekhar, J.; Spitznagel, G.W.; Schleyer, P.v.R. *J. Comput. Chem.* **1983**, 4, 294. (p) Schleyer, P.v.R.; Apeloig, Y.; Arad, D.; Luke, B.T.; Pople, J.A. *Chem. Phys. Lett.* **1983**, *95*, 477. (q) Del Bene, J.E.; Frisch, M. J.; Raghavachari, K.; Pople, J.A. *J. Phys. Chem.* **1982**, *86*, 1529. (r) Lien, M.H.; Hopkinson, A.C. *J. Mol. Struct. Theochem* **1985**, *121*, 1. (s) Taagepera, M.; Summerhays, K.D.; Hehre, W.J.; Topsom, R.D.; Pross, A; Radom, L.; Taft, R.W. *J. Org. Chem.* **1981**, *46*, 891. (t) Eades, R.A.; Scanlon, K.; Ellenberger, M.R.; Dixon, D.A.; Marynick, D.S. *J. Phys. Chem.* **1980**, *84*, 2840. (u) Köhler,

H.-J.; Lischka, H *J. Am. Chem. Soc.* 1979, *101*, 3479. (v) Ikuta, S.; Kebarle; P. *Can. J. Chem.* 1983, *61*, 98. (w) Hopkinson, A.C.; Lien, M.H.; Yates, K.; Mezey, P.G.; Csizmadia, I.G. *J. Chem. Phys.* 1977, *67*, 517. (x) Köhler, H.-J.; Lischka, H. *Theor. Chim. Acta* 1979, *54*, 23. (y) Green, S.; Schor, H.; Siegbahn, P.E.M.; Thaddeus, P. *Chem. Phys.* 1976, *17*, 479. (z) Diercksen, G.H.F.; Kraemer, W.P.; Roos, B.O. *Theor. Chim. Acta* 1975, *36*, 249. (aa) Hudson, R.F.; Eisenstein, O.; Anh, N.T. *Tetrahedron* 1975, *31*, 751. (bb) Hopkinson, A.C.; Csizmadia, I.G. *J. Chem. Soc., Chem. Commun.* 1971, 1291. (cc) Ellenberger, M. R.; Eades,R.A.; Thomsen, M. W.; Farneth, W.E.; Dixon, D.A. *J. Am. Chem. Soc.* 1979, *101*, 7151. (dd) Eades, R.A.; Weil, D.A.; Ellenberger, M.R.; Farneth, W.E.; Dixon, D.A.; Douglass, C.H., Jr. *Ibid.* 1981, *103*, 5372. (ee) Del Bene, J.E.; Radovick, S. *Ibid.* 1978, *100*, 6936. (ff) Umeyama, H.; Morokuma, K. *Ibid.* 1976, *98*, 4400. (gg) Kollman, P.; Rothenberg, S. *Ibid.* 1977, *99*, 1333. (hh) Köhler, H.-J.; Lischka, H. *Chem. Phys. Lett.* 1978, *58*, 175. (ii) Redfern, P.; Scheiner, S. *J. Comput. Chem.* 1985, *6*, 168. (jj) Smith, S. F.; Chandrasekhar, J.; Jorgensen, W.L. *J. Phys. Chem.* 1982, *86*, 3308. (kk) Eades, R.A.; Weil, D.A.; Dixon, D.A.; Douglass, C.H. *Ibid.* 1981, *85*, 981. (ll) Huber, H.; Vogt, J. *Chem. Phys.* 1982, *64*, 399. (mm) Boerth, D.W.; Streitwieser, A., Jr. *J. Am. Chem. Soc.* 1978, *100*, 750. (nn) Del Bene, J.E. *J. Comput. Chem.* 1985, *6*, 296. (oo) Ikuta, S.; Kebarle, P. *Can. J. Chem.* 1983, *61*, 97. (pp) Catalán, J.; de

Paz, J.L.G.; Yanez, M; Elguero, J. *J. Am. Chem. Soc.* 1984, *106*, 6552. (qq) Dill, J.D.; Greenberg, A.; Liebman, J.F. *Ibid.* 1979, *101*, 6814. (rr) Bernardi, F.; Csizmadia, I.G.; Schlegel, H.B.; Wolfe, S. *Can. J. Chem.* 1975, *53* 1144. (ss) Daudel, R.; Kozmutza, C.; Kapuy, E. *Chem. Phys. Lett.* 1975, *36*, 555. (tt) Hopkinson, A.C.; Holbrook, N.K.; Yates, K.; Csizmadia, I.G. *J. Chem. Phys.* 1968, *49*, 3596. (uu) Hopkinson, A.C.; Yates, K.; Csizmadia, I.G. *Ibid.* 1970, *52*, 1784. (vv) Spitznagel, G. W.; Clark, T.; Chandrasekhar, J.; Schleyer, P. v. R. *J. Comput. Chem.* 1982 *3*, 363. (ww) Jasien, P. G.; Stevens, W. J. *J. Chem. Phys.* 1985, *83*, 2984. (xx) Hehre, W. J.; Taagepera, M.; Taft, R. W.; Topsom, R. D. *J. Am. Chem. Soc.* 1981, *103*, 1344. (yy) Reynolds, W. F.; Modro, T. A.; Mezey, P. G.; Skorupowa, E.; Maron, A. *Can. J. Chem.* 1980, *58*, 412. (zz) Pross, A.; Radom, L; Taft, R. W. *J. Org. Chem.* 1980, *45*, 818. (aaa) McKelvey, J.M.; Alexandratos, S.; Streitwieser, A., Jr.; Abboud, J.-L. M.; Hehre, W. J. *J. Am. Chem. Soc.* 1976 *98*, 244. (bbb) Heyne, E.; Raabe, G.; Fleischhauer, J. *Z. Naturforsch., A: Phys., Phys. Chem., Kosmophys.* 1984, *39A*, 593. (ccc) Del Bene, J. E. *Ibid.* 1978, *100*, 1673. (ddd) Kollman, P.; McKelvey, J.; Gund, P. *Ibid.* 1975, *97*, 1640. (eee) Williams, J.E., Jr.; Streitwieser, A., Jr. *Ibid.* 1975, *97*, 2634. (fff) Hehre, W. J.; McIver, R.T., Jr.; Pople, J.A.; Schleyer, P. v. R. *Ibid.* 1974, *96*, 7162. (ggg) Del Bene, J.E.; Vaccaro, A. *Ibid.* 1976, *98*, 7526. (hhh) Hehre, W. J.; Pople, J. A. *J. Am. Chem. Soc.*, 1972, *94*, 6901. (iii)

Ermler, W. C.; Mulliken, R. S.; Clementi, E. *Ibid.* **1976**, *98*, 388. (jjj) Houriet, R.; Rüfenacht, H. Carrupt, P.-A.; Vogel, P.; Tichý, M. *Ibid.* **1983**, *105*, 3417. (kkk) Hoyland, J. R.; Lampe, F. W. *J. Chem. Phys.* **1962**, *37*, 1066. (lll) Hariharan, P. C.; Lathan, W. A.; Pople, J. A. *Chem. Phys. Lett.* **1972**, *14*, 385. (mmm) Scarlett, M.; Taylor, P. R. *Chem. Phys.* **1986**, *101*, 17. (nnn) Dyczmons, V.; Staemmler, V.; Kutzelnigg, W. *Chem. Phys. Lett.* **1970**, *5*, 361. (ooo) Pross, A.; DeFrees, D.J.; Levi, B.A.; Pollack, S. K.; Radom, L.; Hehre, W. J. *J. Org. Chem.* **1981**, *46*, 1693. (ppp) Hopkinson, A. C.; Lien, M. H. *Int. J. Quantum Chem.* **1980**, *18*, 1371. (qqq) Koppel, I. A.; Moelder, U. H.; Palm, V. A. *Org. React. (Tartu)* **1985**, *21*, 3.

9. DPEs and PAs are frequently approximated by the difference between the total energies of the neutral and deprotonated or protonated molecules in *ab initio* studies.

10. (a) Dewar, M. J. S.; *J. Phys. Chem.* **1985**, *89*, 2145. (b) Dewar, M. J. S.; Storch, D. M. *J. Am. Chem. Soc.* **1985**, *107*, 3898.

11. Olivella, S.; Urpi, F.; Vilarrasa, J. *J. Comput. Chem.* **1984**, *5*, 230.

12. Available from Quantum Chemistry Program Exchange (QCPE).

13. (a) Fletcher, R.; Powell, M. J. D. *Comput. J.* **1963**, *6*, 163. (b) Davidon, W.C. *Comput. J.* **1968**, *10*, 406.

14. Komornicki, A.; McIver, J. W., Jr. *J. Am. Chem. Soc.* **1972**, *94*, 2625.

15. Stull, D. R.; Prophet, J. *JANAF Thermochemical Tables* NSRDS-NBS37, **1971**.

16. In cases where two experimental values are quoted for the PA or DPE of a molecule, the average of the two values are used for calculating average unsigned errors.

17. Raghavachari, K.; Whiteside, R.A.; Pople, J.A.; Schleyer, P.v.R., *J. Am. Chem. Soc.* **1981**, *103*, 5649 and references therein.

18. Dyczmons, V.; Kutzelnigg, W. *Theor. Chim. Acta* **1974**, *33*, 239.

19. The C_{3v} structure was not included in the force constant analysis of Ref. 17.

20. Hiraoka, K.; Kebarle, P. *J. Am. Chem. Soc.* **1976**, *98*, 6119.

21. Bischof, P. K.; Dewar, M. J. S. *J. Am. Chem. Soc.* **1975**, *97*, 2278.

22 (a). Lau, Y. K.; Nishizawa, K.; Brown, R.S.; Kebarle, P. *J. Am. Chem. Soc.* **1981**, *103*, 6291. (b) Cavell, R. G.; Allison, D. A. *Ibid.* **1977**, *99*, 4203. (c) Martinsen, D. P.; Buttrill, S.E., Jr. *Org. Mass Spectrom.* **1976**, *11*, 762.

23(a). Pollack, S. K.; Devlin, J.L., III; Summerhays, K.D.; Taft, R.W.; Hehre, W. J. *J. Am. Chem. Soc.* **1977**, *99*, 4583. (b) Catalán, J.; Yáñez, M. *J. Chem. Soc., Perkin Trans. 2* **1979**, 741. (c) Catalán, J.; Yáñez, M. *Ibid.* **1979**, 1627.

24. Alder, R.W.; Sessions, R.B. in *The Chemistry of Amino, Nitroso and Nitro Compounds and Their Derivatives*, Patai, S., Ed.; John Wiley & Sons: Chichester, **1982**; Vol. 2; pp. 785-9.

25. Wolfe, S. *Acct. Chem. Res.* **1972**, *5*, 102.

26. After this manuscript had been submitted, an *ab initio* study[27] of 1,2-diaminoethane appeared in which the hydrogen bonded gauche conformer was found to be 1.28 kcal/mol more stable than the conformer predicted by AM1. This work was, however, carried out using the 4-21G model. Reexamination at a higher level of theory would be of interest.

27. Van Alsenoy, C.; Siam, K.; Ewbank, J. D.; Schäfer, L. *J. Mol. Struct. Theochem* **1986**, *136*, 77.

28. Meot-Ner (Mautner), M.; Hamlet, P.; Hunter, E.P.; Field, F.H. *J. Am. Chem. Soc.* **1980**, *102*, 6393.

29. Buschek, J.M.; Jorgenson, F. S.; Brown, R.S. *J. Am. Chem. Soc.* **1982**, *104*, 5019.

30. Yamdagni, R.; Kebarle, P. *J. Am. Chem. Soc.* 1973, *95*, 3504.

31. Meot-Ner[28] has discussed the discrepancies between his values and those of Yamdagni and Kebarle[30]. The third set of values were derived indirectly from experiment. It may be noted that Yamdagni and Kebarle obtained identical proton affinities (±0.005 kcal/mol) for the α, ω-diamines from propane, pentane and heptane, and also almost identical values (20.6, 20.0, 20.0 e.u.) for the entropies of cyclization. This would certainly not be expected in view of the known steric problems in medium sized rings.

32. Buschek *et al* indicate hydrogen bonding between the hydroxyl hydrogen and nitrogen in the neutral molecule.[29] AM1 optimization of the neutral structure leads to hydrogen bonding between an amine hydrogen and oxygen.

33. A value of -1.9 kcal/mol was reported previously based on results for ethane, propane, n-butane and n-pentane[5].

34. These values are based on the AM1 results for isobutane, isopropylamine and 2-propanol[5], and additional unpublished results for 2-methylbutane, 2-aminobutane and 2-butanol.

35. Catalán, J.; dePaz, J. L. G.; Yañez, M. *J. Mol. Struct. Theochem* **1984**, *107*, 257.

36. (a) Lias, S. G.; Liebman, J. F.; Levin, R. D. *J. Phys. Chem. Ref. Data* **1984**, *13*, 695. (b) Lias, S. G.; Levin, R. D. Unpublished update to Ref. 32a.

37. McMahon, T. B.; Kebarle, P. *J. Am. Chem. Soc.* **1985**, *107*, 2612.

38. Pedley, J.B.; Rylance, G. *Sussex-N.P.L. Computer Analysed Thermochemical Data: Organic and Organometallic Compounds*, Sussex University, 1977.

39. Wagman, D.D.; Evans, W.H.; Parker, V.B.; Schumm, R.H.; Halow, I.; Bailey,S.M.; Churney, K.L.; Juttall, R.L. *The NBS Tables of Chemical Thermodynamic Properties: Selected Values for Inorganic and C_1 and C_2 Organic Substances in S.I. Units J. Phys. Chem. Ref. Data* **1982**, *11*, Suppl. 2.

40. The allenyl anion is not a stationary point on the AM1 surface. The AM1 ΔH_f value was calculated by fixing the <C-C-H value at the value optimized in Ref. 8w and forcing C_s symmetry. The same method was used in Ref. 11 to calculate the MNDO ΔH_f value

(Olivella, S. Personal communication).

41. While ion cyclotron resonance data point to initial formation of acetylenic and propargylic (or allenic) anions, equilibration leads to the dominance of acetylenic anions at long times[42], implying the propynyl anion is the most stable of the three. Likewise the allenyl anion has been found to be more stable than propargyl anion both experimentally[43] and by *ab initio* calculations[8w,44,45].

42. Bartmess, J.E.; Scott, J.A.; McIver, R.T., Jr. *J. Am. Chem. Soc.* 1979, *101*, 6046.

43. Oakes, J. M.; Ellison, G. B. *J. Am. Chem. Soc.* 1983, *105*, 2969.

44. Bushby, R. J.; Patterson, A. S.; Ferber, G. J.; Duke, A. J.; Whitham, G. H. *J. Chem. Soc., Perkin Trans.2* 1978, 807.

45. Wilmshurst, J. K.; Dykstra, C.E. *J. Am. Chem. Soc.* 1980, *102*, 4668. The "allenyl"-like structure optimized in this study had ∠C-C-C=175.9°.

46. Bartmess, J.E.; Burnham. R.D. *J. Org. Chem.* 1984, *49*, 1382.

47. Direct methyl and phenyl substitution on HO^-, H_2N^-, CN^- and HCC^- is not included in calculating these averages due to the excessive errors in the calculated ΔH_f values for these anions.

48. Bartmess, J.E.; Caldwell, G.; Rozenboom, M.D. *J. Am. Chem. Soc.* 1983, *105*, 340.

49. Bartmess, J.E. Personal communication.

50. Taft, R. W. in *Progress in Physical Organic Chemistry*, Taft, R. W., Ed., John Wiley & Sons: New York, 1983; Vol. 14, pp. 276-283.

51. Recent experimental evidence has indicated alkyl group size - acidity relationships are also present in the neutral alcohol molecules. (Chauvel, J.P., Jr.; True, N.S. *Chem. Phys.* 1985, *95*, 435.) This calls into question the relative importance of charge-induced dipole stabilization of anions in determining the relative acidities of alcohols.

52. Bartmess, J.E., intended for publication in *J. Phys. Chem. Ref. Data.*

53. Benson, S. W. *Thermochemical Kinetics, 2nd Ed.*; John Wiley & Sons: New York, 1976; Chapter 2.

54. These values are based on a proposed change to the anchor point at the methanol end of the acidity scale. (Moylan, C.R.; Brauman, J.I. *J. Phys. Chem.* 1984, *88*, 3175.) Moylan et al suggest the discrepancy is the result of difficulties in studying reactions of HF, the previous anchor point. Bartmess has indicated[52] the discrepancy may be due to the temperature correction problem in ICR work[1b,36a].

55. Wiberg[56], and Ibrahim and Schleyer[57], have shown that good estimates of heats of formation of organic molecules can be obtained from *ab initio* total energies (6-31G* basis set) by applying empirical corrections based on group additivity relationships or their equivalent. A similar scheme could probably be developed for calculating PAs and DPEs. Such an approach could not, however, be used to study reactions for the same reason that molecular mechanics cannot be used in this connection, i.e. the fact that empirical corrections cannot be derived for the variable partially bonded groups present in transition states.

Koppel *et al*[8qqq] have recently carried out linear regressions on *ab initio* PAs and DPEs calculated using various basis sets (STO-3G, 3-21G, 4-31G, 6-31G*, 3-21+G, 4-31+G) and with the 6-31G** basis set using fourth order Moeller-Plesset perturbation theory. This procedure leads to a much better fit with experiment. However, even with these corrections, the results become markedly superior to our uncorrected ones only at or above the 6-31G* level.

The computing time required limits such procedures to reactions of very small molecules.

56. Wiberg, K. B. *J. Comput. Chem* 1984, *5*,

57. Ibrahim, M.R.; Schleyer, P. v. R. *J. Comput. Chem* 1985, *6*, 157.

Chapter 4

An AM1 and MNDO Study of the Condensation Reaction of Polyketide Biosynthesis

Introduction

Polyketide biosynthesis is responsible for a diverse group of natural products. Defined in biosynthetic terms as opposed to structural terms, polyketides, in general, comprise structures derived essentially from poly-β-ketomethylene chains, $-[CHRCO]_n-$, R commonly, but not exclusively, hydrogen.[1] As such, this category of compounds includes fatty acids, as well as compounds with more exotic structures, such as 6-methylsalicylic acid.[2]

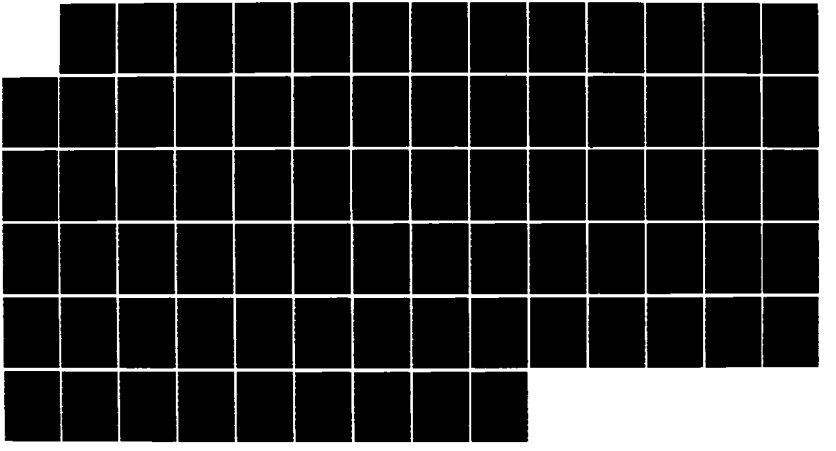
While the ultimate products vary greatly, the biosynthesis of fatty acids and other polyketides follow an essentially similar pathway[3,4] involving sequential condensation of two-carbon subunits onto a lengthening carbon skeleton. The diversity of the products arises from variations in the reactions occurring between successive condensations.

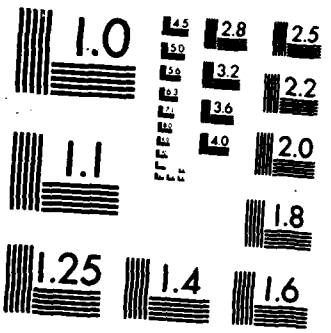
The two-carbon subunits could be derived directly from simple aliphatic acids:

$-CH_2CO-$ 'acetate'

AD-A173 032 APPLICATION OF THE AM1 AND MNDO SEMIEMPIRICAL QUANTUM MECHANICAL MOLECULAR MODELS(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH K M DIETER DEC 86 2/2

UNCLASSIFIED AFIT/CI/NR-86-183D F/G 7/4 NL



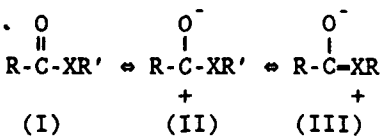


$-CH(CH_3)CO-$ 'propionate'

$-CH(C_2H_5)CO-$ 'butyrate'

In actuality, they enter the synthetic process as malonic acid derivatives (HOOCCHRCOOH), which are more effective nucleophiles.[1] While many natural polyketides can be described as "wholly acetate-derived", having been synthesized from an initial acetyl "primer" unit by successive condensations of malonate groups (which are synthesized by carboxylation of acetyl units[4]) with coincidental evolution of CO_2, some enzymes are known to accept different acyl "primer" units and substituted malonate groups.[1]

With the exception of the transacylase activities, which transfer acetyl and malonyl groups between coenzyme A (CoA) and the enzymes, thiol esters are used in place of alcohol esters.[5] Acyl and malonyl groups are stronger nucleophiles and electrophiles in thiol esters than in alcohol esters. According to the resonance explanation normally given for this difference in behavior,[4,6] nonbonding electrons of the alcohol oxygen atom of alcohol esters are delocalized into the carbonyl group, and all three resonance structures shown below are important. In thiol esters, the overlap



(I) (II) (III)

X=O,S

of the 3p orbitals with the π system of the carbonyl group is less effective, and resonance structure (III) is less important. Consequently, the carbonyl carbon is more electrophilic in thiol esters than alcohol esters, and the thiol esters are better acylating agents. While there is some evidence against this argument,[7] the increased activity in thiol esters is not disputed. Indeed, this increased activity was evident in the results of a gas phase study by Bartmess *et al*,[8] in which acetaldehyde enolate reacted with methyl thiolformate to form Claisen condensation product, while only a trace of this product was found when methyl formate was used as the electrophile. In addition to the greater electrophilicity of the carbonyl carbon, α-hydrogens are more acidic in thiol esters due to more effective delocalization of the negative charge of the conjugate base. Hence, thiol esters also exhibit more nucleophilic activity than alcohol esters.[4,6]

Given, then, the thiol ester form of the substrates and their associated nucleophilic and electrophilic activity, the simplest model for the condensation in polyketide biosynthesis involves formation of an anion and subsequent Claisen type condensation.[6] This proposal was rejected by Arnstadt *et al*, however, on the basis of isotopic studies.[9] In their studies, dideutero-malonyl-CoA was used as the substrate for yeast fatty acid synthetase (FAS), and no primary isotope effect on the rate of fatty acid synthesis was observed. Additionally, when they investigated the condensation reaction separately from other FAS reactions by

using the β-ketoacyl-acyl-carrier-protein synthetase of *Esterichia coli* (see below) in the presence of tritiated water, they did not observe any incorporation of tritium in the acetoacetyl-thiol ester product. From these results, Arnstadt concluded free carbanions are not involved in the condensation, and the formation of the new carbon bond is coupled with the cleavage of the carbonyl bond of the malonyl group.

These conclusions are based on a model in which enzyme catalyzed reactions take place in an analogous manner to reactions in solution. Dewar and Storch recently introduced an alternative model in which solvent molecules are excluded from the active sites of enzymes. As a result, the enzyme catalyzed reactions occur in a solvent free environment, similar to gas phase reactions.[10] If this is the situation for the condensation reaction of polyketide biosynthesis, the participation of a free carbanion cannot be eliminated by the results of Arnstadt *et al*,[9] and the decarboxylation of the malonate thiol ester and condensation with the acyl thiol ester may involve carbanions and may not be concerted. Consequently, it was of interest to study the decarboxylation of malonate thiol esters and malonate alcohol esters, and their condensation with acetyl thiol and alcohol esters, using the semi-empirical quantum mechanical models MNDO[11] and AM1[12]. Since these models are parameterized using gas phase experimental data, results should reflect what occurs if the solvent molecules are excluded from the active site of the enzyme.

Procedure

The standard closed shell MNDO[11,13] and AM1[12] procedures as implemented in the AMPAC[14] package of computer programs were used for all calculations. With the exception of complexes in which separation of charged groups was required (see below), all geometries were optimized by minimizing the energy with respect to all geometrical variables using the DFP method.[15] Minima and transition states were characterized whenever possible by insuring the applicable Hessian (second derivative) matrices had *exactly* zero or one negative eigenvalue, respectively.[16,17] Entropies were calculated[19] when it was anticipated the entropy of activation and the entropy of reaction would be significant. The entropy contribution for an internal rotation lost during the course of the reaction was calculated using the tables of Pitzer and Gwinn.[20]

Model

The mechanism for fatty acid synthesis is the most thoroughly studied of the polyketide synthesis mechanisms.[1] Since it is anticipated the syntheses of other polyketides, and specifically the condensation portion of the syntheses, follow a pathway essentially similar to that of fatty acid synthesis,[4,6] discussion is limited to fatty acid synthesis.

Fatty acid synthetases are found in two general forms. In *Esterichia coli* and several plants, fatty acid synthesis is catalyzed by at least six enzymes and acyl carrier protein (ACP), nonassociated or loosely associated in a complex and separable from each other using conventional techniques for enzyme fractionation and purification.[5,21,22] FAS from yeast and animals, on the other hand, are multifunctional molecular enzyme complexes.[5,22,23] Most work on elucidating structural features of the active site of the condensing activity has been done on the associated type of synthetase, and discussion is further restricted to this type.

The condensation reaction involves two active thiol groups.[24] One thiol group belongs to a cysteine residue localized in the condensing activity of the FAS (hereafter referred to as cys-SH). The second thiol group belongs to the 4'-phosphopantetheine group of the ACP of the FAS (hereafter referred to as pant-SH). These active thiol groups can approach within approximately 2 Å of each other.[24,25] Prior to condensation, the acetyl transacylase activity of the FAS transfers an acetyl "primer" group from acetyl coenzyme A (acetyl-CoA, CH_3COS-CoA) to the pant-SH, from which it is transferred again, this time to the cys-SH[24,26]:

$$CH_3\overset{O}{\overset{\|}{C}}S\text{-CoA} + \text{pant-SH} \rightarrow CH_3\overset{O}{\overset{\|}{C}}S\text{-pant} + \text{CoA-SH}$$

$$CH_3\overset{O}{\overset{\|}{C}}S\text{-pant} + \text{cys-SH} \rightarrow CH_3\overset{O}{\overset{\|}{C}}S\text{-cys} + \text{pant-SH}$$

The malonyl transacylase activity then transfers a malonate group from malonyl-coenzyme A (malonyl-CoA, HO_2CCH_2COS-CoA) to the pant-SH[24]:

$$HO_2CCH_2\overset{O}{\overset{\|}{C}}S\text{-CoA} + \text{pant-SH} \rightarrow HO_2CCH_2\overset{O}{\overset{\|}{C}}S\text{-pant} + \text{CoA-SH}$$

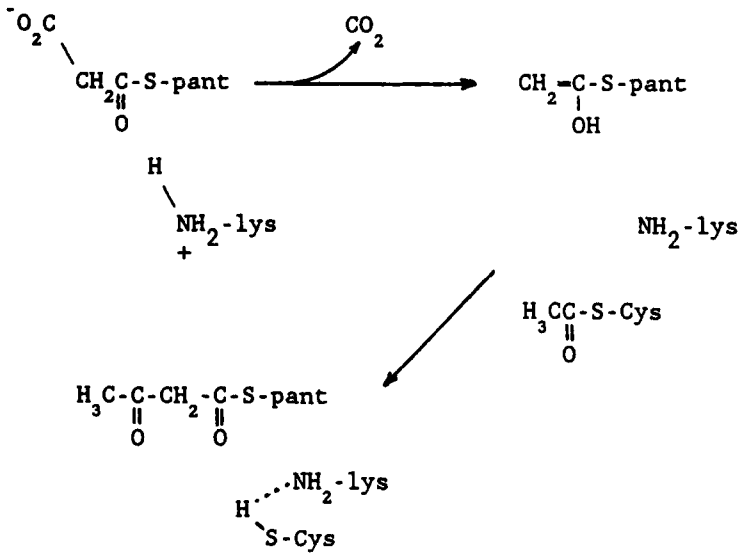
The mechanisms of these transacylases are not known.[5] Decarboxylation of the malonate group and condensation with the acetyl group, with inversion of configuration,[27] follows:

$$CH_3\overset{O}{\overset{\|}{C}}S\text{-cys} + HO_2CCH_2\overset{O}{\overset{\|}{C}}S\text{-pant} \rightarrow$$

$$CH_3\overset{O}{\overset{\|}{C}}CH_2\overset{O}{\overset{\|}{C}}S\text{-pant} + \text{cys-SH} + CO_2$$

with decarboxylation providing the ultimate driving force for the reaction.[1] The β-ketoacyl thiol ester, still bonded to ACP, then undergoes reduction, dehydration and reduction, again,[24] to the saturated acyl thiol ester. After transfer of the acyl thiol ester to the cys-SH, another malonyl group is transferred from malonyl-CoA to the pant-SH and the condensation-reduction-dehydration-reduction cycle begins again.

The ϵ-amino group of a lysine residue, as well as a second cysteine thiol group, are also present in the active site of the condensing activity.[28] The ϵ-amino group is thought to perturb the pK of the active cys-SH, with the two groups existing as either a hydrogen-bonded complex between the cys-SH and the lys-NH_2 groups,[29] an ionic complex between cys-S^- and lys-NH_3^+,[24,28] or a hybrid of the two, in the absence of an acyl group on the active cysteine. A mechanistic role for the lys-NH_3^+ group in the decarboxylation of HO_2CCH_2COS-pant as a proton donor/electron sink to stabilize the incipient enolate anion is also possible.[24,28] The catalytic role,



if any, of the second cysteine thiol group and the significance of its close proximity to the active thiol groups are unknown. However, since it is present in yeast FAS as well as chicken liver FAS, it also may play a role in the condensation.[28]

In the calculations, a CH_3S^- group was used to model the pantetheine and cysteine functions of the acetyl and malonyl thiol esters. Since AM1 is able to reproduce hydrogen bonds whereas MNDO is not,[11,12,30] and sulfur has not, as of yet, been parameterized for AM1, corresponding alcohol esters were used in complexes involving hydrogen bonding. The ϵ-NH_2/NH_3^+ group of the lysine residue was modeled by methyl amine/ammonium ion rather than ammonia/ammonium ion. This allowed internal coordinates of the methyl carbon to be fixed while still allowing greater freedom in the position of the $-NH_3^+$ group with respect to the malonyl alcohol ester. The fixing of internal coordinates of the methyl carbon of methyl ammonium ion, and of the carboxylate carbon of the malonyl alcohol ester, was required in some calculations to prevent the negatively-charged carboxylate group and the positively-charged ammonium group from collapsing.[31] In order to calculate entropic contributions to the decarboxylation, the required Hessian matrix had to be calculated without constraining any internal coordinates. For this purpose formic acid was used as the proton donor for the hydrogen bond to the carbonyl group of the malonyl alcohol ester. A formic acid molecule was also used to model the second cys-SH. In this situation formic acid was chosen because a possible role of the second cys-SH may be in interacting with the carbonyl oxygen of the acyl group bonded to the active cys-SH, and the AM1 calculated deprotonation enthalpy of formic acid, 354 kcal/mol,[30] is close to the experimental value for methane thiol, 359 kcal/mol.[32]

Table 1. Comparison with experiment of MNDO and AM1 calculated heats of formation (kcal/mol).

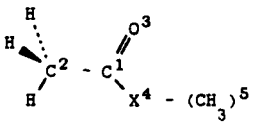
Molecule	Expt'l	MNDO	Error	AM1	Error
CH_3COSCH_3	-47.8[a]	-51.0	-3.2		
$^-CH_2COSCH_3$		-54.6			
$CH_3CO_2CH_3$	-98.4[b]	-93.6[c]	4.8	-96.4[d]	2.0
$^-CH_2CO_2CH_3$	-94.6[e]	-90.4	4.2	-94.2[f]	0.4
$HO_2CCH_2COSCH_3$	-136.2[a]	-135.7	0.5		
$^-O_2CCH_2COSCH_3$		-154.5	(10.5)[k]		
$HO_2CCH_2CO_2CH_3$	-186.7[a]	-177.7	9.0	-182.8	3.9
$^-O_2CCH_2CO_2CH_3$		-192.3	(19.0)[k]	-207.4	(9.9)[k]
CH_3NH_2	-5.5[b]	-7.5[c]	-2.0	-7.4[d]	-1.9
$CH_3NH_3^+$	147.6[g]	161.8[h]	14.2	148.7[f]	1.1
HCO_2H	-90.5[b]	-92.7[c]	-2.2	-97.4[d]	-6.9
HCO_2^-	-112.5[e]	-101.7[i]	10.8	-110.0[f]	2.5
CO_2	-94.1[j]	-75.4[c]	18.7	-79.8[d]	14.3

[a] ΔH_f estimated using Benson's group additivity method.[34] The contribution for S(CO)(C) was estimated as S(CO)(H) + $(S(C)_2$-S(C)(H)) + $(S(C)(C_d)$-$S(C)_2)$. [b] See ref. 35. [c] See ref. 11. [d] See ref. 12. [e] ΔH_f calculated using the experimental values of the DPE[32] and ΔH_f[32] of the protonated form, and the experimental value for $\Delta H_f(H^+)$.[36] [f] See ref. 30. [g] ΔH_f calculated using the experimental values for the proton affinity[37] and ΔH_f[32] of the deprotonated form, and the experimental value for $\Delta H_f(H^+)$.[36] [h] See ref. 33. [i] See ref. 38. [j] See ref. 39. [k] Estimated using the known tendency of MNDO and AM1 to overestimate the DPEs of aliphatic carboxylic acids by 10 kcal/mol and 6 kcal/mol, respectively.[30,33]

Results

A. Calculation of Model Compounds. Recent studies have indicated that, with few exceptions, AM1 and MNDO give good estimates of proton affinities (PA) and deprotonation enthalpies (DPE) of neutral molecules.[30,33] The errors in PAs and DPEs, as well as in the heats of formation of the corresponding cationic acids and anionic bases, are similar in magnitude to those in the heats of formation of the neutral molecules. Results are presented in Table 1 for the calculated heats of formation of the model compounds in this study. Errors are generally within the average errors for MNDO[11] and AM1[12], with the performance of AM1 generally better than MNDO. Three previously indicated[11,33] problem areas are evident for MNDO: the conjugate acid of methyl amine, the conjugate base of formic acid and carbon dioxide. The lone significant discrepancy for AM1 is carbon dioxide, which was also previously indicated.[12] Since these errors are known, they can be allowed for when analyzing the results of reaction path calculations. Most significant in this respect are the errors in the conjugate bases of carboxylic acids. MNDO and AM1 are known to overestimate the DPEs of aliphatic carboxylic acids by approximately 10 kcal/mol and 6 kcal/mol, respectively.[30,33] These values, then, are used to estimate errors for the malonate derivatives in Table 1.

Table 2. Calculated atomic charges for methyl acetate, methyl thiolacetate and their conjugate bases resulting from loss of an α-H^+.



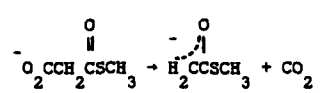
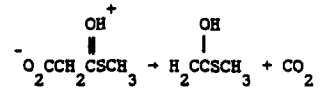
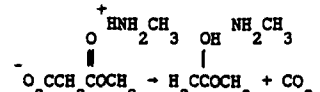
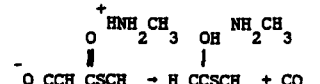
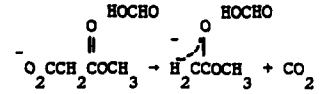
	AM1			MNDO		
	Neutral	Anion	Change	Neutral	Anion	Change
X=O						
C^1	0.30	0.34	0.04	0.35	0.38	0.03
C^2	-0.22	-0.64	-0.42	0.05	-0.50	-0.55
O^3	-0.35	-0.58	-0.23	-0.36	-0.61	-0.25
O^4	-0.28	-0.35	-0.07	-0.35	-0.39	-0.04
$(CH_3)^5$	0.21	0.07	-0.14	0.23	0.11	-0.12
X=S						
C^1				0.14	0.22	0.08
C^2				0.01	-0.46	-0.47
O^3				-0.29	-0.56	-0.27
S^4				0.03	-0.16	-0.19
$(CH_3)^5$				0.03	-0.07	-0.10

The MNDO DPE[40] for methyl thiolacetate (363.6 kcal/mol) is significantly lower than the MNDO, AM1, and experimental DPEs for methyl acetate (370.4,[40] 369.4[30,40] and 371.0[32] kcal/mol, respectively). This is in agreement with the trend expected, considering the argument concerning more effective charge delocalization in the enthiolate anion (see above). However, calculated atomic charges (Table 2) do not indicate a significant difference in the amount of the anionic charge absorbed by the carbonyl oxygen in methyl acetenthiolate anion as compared to methyl acetenolate anion. A larger change is seen in the amount of charge

absorbed by the sulfur atom as opposed to the alcohol oxygen atom. The calculations, then, model the difference in electronic behavior between the thiol ester and the alcohol ester more as a polarization effect than as a resonance effect. Comparison of the MNDO calculated heat of formation of CH_3S^- (-13.8 kcal/mol) with the experimental value (-12.2 kcal/mol[32]) does not indicate a significant overstabilization of negative charge by sulfur in MNDO calculations. However, contrary to conclusions based on ^{13}C NMR results,[7] the carbonyl carbon is predicted to carry a lower positive charge in the thiol ester than in the alcohol ester. Hence, polarization effects of the sulfur atom are probably overestimated in the calculations, while resonance effects are underestimated.

B. **Decarboxylation.** Thermodynamic results for various decarboxylation schemes are given in Table 3. The decarboxylation of $^-O_2CCH_2COSCH_3$ was simulated by treating the $O_2C\cdots CH_2R$ bon the reaction coordinate, and optimizing all other geometric variables while incrementally increasing the value of the reaction coordinate. The reaction was predicted to be endothermic by 24.5 kcal/mol, with a forward activation barrier of 29.0 kcal/mol. The vibration corresponding to the negative eigenvalue of the Hessian matrix of the transition state was the $O_2C\cdots CH_2R$ stretch. The value of the reaction coordinate was 2.41 Å in the transition state.

Table 3. Thermodynamic values calculated for decarboxylation schemes treating the $O_2C\cdots CH_2R$ bond length as the reaction coordinate.[a,b]

Reaction	Method	$\Delta H^{\neq}$	$\Delta S^{\neq}$	$\Delta G^{\neq}$	ΔH_{rxn}	ΔS_{rxn}	ΔG_{rxn}
	MNDO	29.0	3.04[c]	28.1	24.5	33.3[c]	14.5
					16.5[d]		6.3[d]
	MNDO	-	-	-	-45.8		-55.8[e]
	AM1	10.7		9.6[e]	-20.9		-30.9[e]
	AM1	10.9[f]					
	MNDO	16.9[f]					
	MNDO	13.3[f]					
	AM1	-	-	-	25.4	30.9	16.0
					20.9[d]		11.6[d]

[a] Enthalpies and free energies are in kcal/mol; entropies are in e.u. [b] T=300K. [c] The barrier for the hindered rotation of the carboxylate group in the reactant was estimated as 0.5 kcal/mol, the barrier given by Benson for the CH_3-CO_2H hindered rotation.[3,4] [d] These figures were calculated using the experimental $\Delta H_f(CO_2)$ and corrected heats of formations for carboxylate ions based on the errors given in Table 1. [e] The change in entropy was assumed to be the same as that calculated for decarboxylation of the malonate thiol ester. [f] The O···H hydrogen bond length was restricted to that optimized by AM1 in the reactant complex.

The large MNDO error in the heat of formation of CO_2 could account for most of the heat of reaction and for part of the activation barrier. However, since MNDO underestimates the stability of the carboxylate anion relative to the carboxylic acid by approximately 10 kcal/mol, the two errors largely offset each other. The result is that the reaction is still predicted to be endothermic by a significant amount. Since these corrections cannot be assumed valid for the heat of formation of the transition state, no corrected value can be estimated for the activation barrier. In any case, it is clear the activation barrier is formidable.

With the gain of three translational and three rotational degrees of freedom at the expense of five vibrations and a hindered internal rotation, the decarboxylation will be favored entropically. The increase in entropy is not large enough, however, to offset the unfavorable enthalpy changes. It is reasonable to conclude, then, that there must be other interactions with the malonate thiol ester to prompt decarboxylation.

As discussed above, the carbonyl group *beta* to the carboxylate group is expected to activate the malonate thiol ester toward decarboxylation by allowing the delocalization of the negative charge of the product enthiolate anion. Jencks has pointed out that the carbonyl group is an even more active electron sink when it is protonated.[41] As expected on this basis, when a proton was added to the carbonyl group and the geometry optimized, CO_2 was

lost without activation. The decarboxylation was exothermic by 45.8 kcal/mol and exergonic by 55.8 kcal/mol.

A logical source of this proton would be an ammonium ion derived from the ϵ-NH_2 group of the lysine residue in the active site. To simulate it, methyl ammonium ion was positioned in the plane of the ester linkage of the malonate alcohol ester, allowing hydrogen bonding between the ammonium ion and the ester function. Bond angles and dihedral angles were frozen to prevent the ammonium ion and the carboxylate group from collapsing. The decarboxylation was now found to be exothermic and exergonic (-20.9 and -30.9 kcal/mol, respectively). An apparent transition state[17] at an $O_2C\cdots CH_2R$ distance of 2.06 Å defined an activation barrier of 10.7 kcal/mol. A proton was transferred from the methyl ammonium ion to the carbonyl group of the ester at a $O_2C\cdots CH_2R$ distance of 3.9 Å. Interestingly, however, when the errors for carboxylate anions and CO_2 were taken into account, the enthalpy of reaction for the decarboxylation was approximately zero prior to proton transfer (not shown in Table 3). Furthermore, when the bond between the proton on the methyl ammonium ion and the carbonyl oxygen was treated as the reaction coordinate, the proton was transferred to the carbonyl ($\Delta H^{\neq}$ = 12.2 kcal/mol; ΔH_{rxn} = 8.2 kcal/mol) *without* resulting in decarboxylation (see below).

This model was oversimplified in three respects. First, it ignored possible electrostatic interactions with the electrophile (thiol ester) with which the nascent carbanion reacts in the next step. This will be addressed later when discussing the possibility of a concerted decarboxylation-condensation process. Secondly,it replaced oxygen for sulfur in the ester linkage. Thirdly, it restricted some of the internal coordinates of the malonate ester - ammonium ion complex.

To evaluate the effect of substituting oxygen for sulfur in the ester linkage, the hydrogen bond length was fixed at the AM1 optimized value during decarboxylation. Decarboxylation was then simulated using AM1 and MNDO with the alcohol ester complex, and using MNDO with the thiol ester complex. Three contributions to changes in the reaction profile were taken into account when evaluating the results: (1) the effect of restricting the hydrogen bond length, (2) the effect of using MNDO instead of AM1, and (3) the effect of having oxygen rather than sulfur in the ester linkage. Restricting the hydrogen bond length increased the barrier only from 10.7 to 10.9 kcal/mol. Using MNDO rather than AM1 increased the barrier another 6.0 kcal/mol. This is consistent with previous comparisons of AM1 and MNDO calculated activation barriers.[12,42] Additionally, the MNDO barrier for the alcohol ester was 3.6 kcal/mol larger than that for the thiol ester. Using these corrections, the barrier to decarboxylation of the malonate thiol ester with an optimized hydrogen bond length can be estimated as 7.1

kcal/mol.

Next considered was the effect of the restrictions on bond angles and dihedral angles of the methyl ammonium ion and the carboxylate group, which were needed to prevent the oppositely charged groups from reacting with one another. To evaluate this, formic acid was used to form a hydrogen bond in place of methyl ammonium ion, and the geometry was fully optimized. Reactants and products of the decarboxylation were characterized as true minima. While the decarboxylation was endothermic, the enthalpy increased smoothly throughout the reaction, with no transition state between reactant and products. It is therefore possible that at least part of the apparent activation barrier in the decarboxylation of the malonate ester - ammonium ion complex, and certainly the corresponding thiol ester complex, is an artifact due to restrictions placed on the geometry in the calculations.

The reaction path calculated for the decarboxylation of the malonate ester - ammonium ion complex ultimately involved proton transfer. When the proton transfer from the methyl ammonium ion to the malonate ester was treated as the reaction coordinate, there was an activation barrier of 12.2 kcal/mol. The question remains, then, as to whether the apparent barrier to decarboxylation might be partly due to the barrier for proton transfer. There is evidence, however, that this is not the case. First, the barrier to decarboxylation occurs much earlier along the reaction profile than

the proton transfer. In the apparent transition state for the decarboxylation, the H-N bond for the proton ultimately transferred to the carbonyl oxygen of the ester has stretched only 0.01 Å from what it was in the reactant complex. Additionally, restricting the length of the O···H hydrogen bond increased the barrier to decarboxylation by only 0.2 kcal/mol (see above). Finally, proton transfer alone did not cause decarboxylation. These results indicate that the barrier to proton transfer plays little or no role in the decarboxylation.[43]

This raises yet another question: Is proton transfer required at all, or is the electrostatic interaction between the malonate thiol ester and the ammonium ion alone sufficient to initiate decarboxylation? When the ammonium ion of the malonate ester - methyl ammonium ion - methyl acetate complex was replaced by a pure ionic charge 1.85 Å from the carbonyl oxygen of the malonate (the same distance as $NH\cdots O{=}C_{malonate}$ just prior to proton transfer during decarboxylation), the enthalpy of activation for decarboxylation was 7.2 kcal/mol. When the separation was increased to 2.25 Å, a distance large enough to inhibit proton transfer, the barrier to decarboxylation increased only to 12.0 kcal/mol. The correction for using the alcohol ester rather than the thiol ester would decrease this barrier to less than 10 kcal/mol. It is feasible, therefore, the charge of the ammonium ion alone would prompt decarboxylation in the enzyme.

As noted above, it has been claimed that free carbanions cannot be involved in the enzyme reaction because of the lack of tritium incorporation when it is carried out in tritiated water.[9] Such exchange will, however, be impossible if water is excluded when the malonate is absorbed into the active site. To check the effect of water, five water molecules were added to the model system. Whereas the protonated alcohol ester and the protonated thiol ester both lost CO_2 without activation in the absence of water, the "hydrated" protonated malonate ester did not dissociate. When the $O_2C\cdots CH_2R$ bond length was treated as the reaction coordinate, a 5.7 kcal/mol barrier to decarboxylation was predicted by AM1. Given that the barrier would be much larger for a mechanism involving interaction of the β-carbonyl group with an ammonium ion, and that water molecules could also weaken this interaction, these results support the suggestion that a crucial factor in enzyme reactions is the exclusion of water from the reacting system.

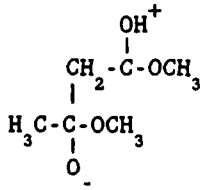
The decarboxylation of the malonate thiol ester in a solvent free environment thus appears to be strongly exothermic and exergonic when an ammonium ion is near enough for proton transfer to occur during the reaction. However, even without proton transfer, the enthalpy of reaction is approximately zero. Since the entropy of decarboxylation is negative, the free energy change for the reaction without proton transfer will be negative. Furthermore, there is a low enthalpic barrier to decarboxylation under these conditions (<10 kcal/mol).

C. **Condensation.** The second step, involving Claisen condensation of the anion and ester, was studied by treating the distance between the methene carbon of the enolate or enthiolate anion and the carbonyl carbon of the corresponding alcohol or thiol ester as the reaction coordinate. As the value of the reaction coordinate was incrementally decreased, all other internal coordinates were fully optimized, except where a few internal coordinates were frozen to prevent the collapsing of oppositely charged groups.[31] Table 4 lists the enthalpy of activation and enthalpy of reaction results. Entropy changes will not play a significant role in the corresponding enzymatic reaction since all intermediates remain bound to the enzyme. Consequently, entropies were not calculated.

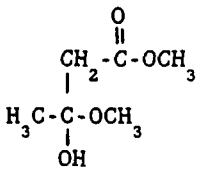
Condensation of the enthiolate anion with methyl thiolacetate was predicted by MNDO to be 2.6 kcal/mol exothermic, with an activation barrier of 15.3 kcal/mol. Given the propensity of MNDO to overestimate activation barriers,[12] the condensation of the oxygen analogs was simulated using MNDO and AM1. MNDO predicted essentially the same barrier, 14.5 kcal/mol. AM1, on the otherhand, predicted the exothermic formation of a charge-dipole complex at 4.5 Å ($\Delta H_{complexation}$ = -7.4 kcal/mol), which was characterized as a true minimum. From the charge-dipole complex there was a 5.8 kcal/mol barrier to formation of the tetrahedral adduct. The lower barrier calculated by AM1 as compared to MNDO is consistent with previous results for the formation of tetrahedral complexes.[42] Since

the positive charge on the carbonyl carbon of the thiol ester is underestimated by MNDO, the barrier to condensation for the sulfur compounds should be at least as low. Additionally, with the enthiolate anion formed while bonded in the active site of the enzyme in a complex tight enough to exclude water molecules, it may be formed closer to the methyl thiolester than the 4.5 Å of the methyl acetate - methyl acetenolate complex. In this case, the barrier to condensation of the enthiolate anion and thiol ester in the enzyme would be less than 5.8 kcal/mol from the point on the reaction coordinate at which the enthiolate anion is formed.

As modelled by AM1, however, the decarboxylation of the malonate ester ultimately involves proton transfer from the methyl ammonium ion to the carbonyl oxygen of the enolate anion. The condensation, in this case, would involve an enol rather than an enolate anion. Without proton transfer simultaneous with the condensation, a zwitterionic product, 1, would be formed:



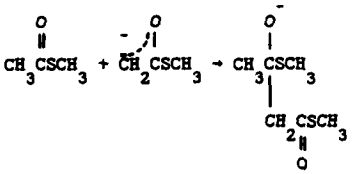
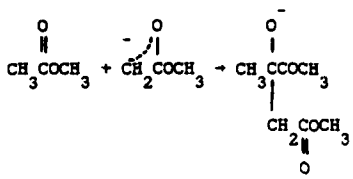
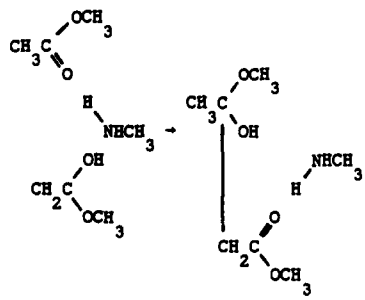
1



2

This product is not a minimum on the AM1 surface, the proton transferring to form the hemiacetal, 2, when the geometry is fully optimized.

Table 4. Thermodynamic values calculated for condensation schemes treating the $RCH_2 \cdots C{=}O$ distance as the reaction coordinate.

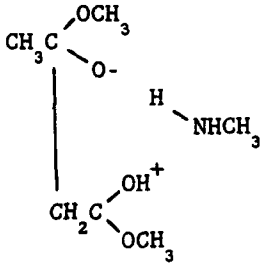
Reaction	Method	$\Delta H^{\neq}$ [a]	ΔH_{rxn} [a]
	MNDO	15.3	-2.6
	MNDO	14.5	
	AM1	-1.6 [b] 5.8 [c]	-9.3 [b] -1.9 [c]
	AM1	- [d]	-26.1 [e]

[a] kcal/mol. [b] Values based on the heats of formation of the isolated reactants. [c] Values based on the heat of formation of the charge-dipole complex between the reactants (see text). [d] A transition state was not found (see text). [e] This value differs by a few kcal/mol depending on the final distance of the amine from the hemiacetal, in this case 2.8 Å from the hydroxyl hydrogen.

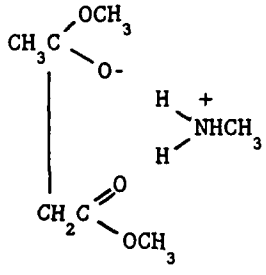
In the enzyme, the proton transfer could be envisaged in either of two ways: (1) transfer directly from the carbonyl oxygen of the enol moiety to the carbonyl oxygen of the acetate moiety via a six membered cyclic transition state, or (2) proton transfer from the enol moiety to the amine and proton transfer from the amine to the methyl acetate moiety. The latter picture is essentially a proton relay model, and is analogous to the "charge transfer mechanism" of chymotrypsin.[46] Both of these processes would give the same net result, a hemiacetal and methyl amine, and are exothermic. A grid search for the direct proton transfer during condensation indicated the enthalpy of activation would be of the order of 30 kcal/mol. Given the size of this barrier and the proximity of the amine group to the condensing moieties in the enzyme, as well as its involvement to this point in the reaction, direct proton transfer is not likely.

To obtain a geometry for the complex involved in the proton-relay model, the amine group was positioned to interact with the carbonyl groups of the malonate ester and methyl acetate. Constraining only the bond angle and dihedral angle of the carboxylate carbon to keep the carboxylate group approximately perpendicular to the plane of the ester linkage,[31] the complex optimized at a separation of 3.68 Å between the two carbons which would form the new C-C bond upon condensation ($N\text{-}H\cdots O{=}C_{malonate}$ 2.12 Å). When the carboxylate group was removed and the geometry was optimized with the carbon-carbon distance fixed at 3.68 Å, a

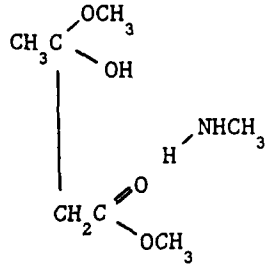
proton transferred from the amine to the carbonyl oxygen of the enolate to form the enol. The carbon-carbon distance was then treated as the reaction coordinate and incrementally decreased while optimizing all other internal coordinates. The enthalpy increased steadily until the apparent intermediate 3 was obtained 30.1 kcal/mol above the initial complex. A force constant analysis



3



4



5

indicated 3 was not a true minimum. Instead, there was one negative eigenvalue corresponding to movement of the proton from the enol moiety to the amine. The resulting complex, 4, was also not a minimum, optimizing instead to 5. The hemiacetal, 5, then, was characterized as the only true minimum resulting from formation of the new carbon-carbon bond between the enol and methyl acetate. Hence, the "true" reaction coordinate involved not only the formation of the carbon-carbon bond, but also a double proton transfer. However, all efforts to find the transition state along this reaction coordinate failed (see below).

D. Concerted Decarboxylation-Condensation. To detect a coupling between these reactions, grid searches were calculated using various geometries of the malonate thiol ester - methyl thiol acetate complex and the oxygen analog, as well as the malonate ester - methyl ammonium ion - methyl acetate complex. The $O_2C\cdots CH_2R$ bond length and the separation between the carbons forming the new bond were treated as the two reaction coordinates. In all cases, decarboxylation occurred essentially independently of condensation. Any path involving significant breaking of the bond to the carboxylate group and simultaneous formation of the new carbon-carbon bond involved an activation barrier of at least 50 kcal/mol.

It seemed possible that the second cysteine in the active site of the condensing enzyme might participate by interacting with the carbonyl group of the acetate, increasing its electrophilicity. However, when the calculation was repeated with an added molecule of formic acid to model this interaction, results were not significantly different.

The presence of methyl acetate did facilitate decarboxylation of the malonate ester. This effect was manifested in two ways. When the $O_2C\cdots CH_2R$ distance was treated as the reaction coordinate in the malonate ester - methyl ammonium ion - methyl acetate complex described above (C···C distance fixed at 3.68 Å), the enthalpy of activation was only 5.2 kcal/mol, as opposed to

10.7 kcal/mol for the malonate ester - methyl ammonium ion complex. Secondly, although there was again no proton transfer from the ammonium ion to the oxygen of the incipient enolate anion until well after the transition state for decarboxylation was passed, when the $NH\cdots O{=}C_{malonate}$ distance was treated as the reaction coordinate, proton transfer did result in decarboxylation. It was possible this effect resulted from the different relative position of the ammonium ion in the malonate ester - methyl ammonium ion - methyl acetate complex as opposed to in the malonate ester - methyl ammonium ion complex, rather than from the presence of methyl acetate. To test this possibility, the bond angle of the nitrogen atom of the ammonium ion relative to the malonate ester was frozen and the methyl acetate was removed from the calculation. The dihedral angle of the nitrogen atom did not change significantly and was allowed to optimize. This time decarboxylation did not occur upon proton transfer. Hence, the facilitation of decarboxylation was the result of the presence of methyl acetate, not the change in the relative position of the ammonium ion. Assuming the same correction as before for using the alcohol ester instead of the thiol ester, and the same value for the entropy of activation for decarboxylation, the free energy of activation for decarboxylation can be estimated as less than 1 kcal/mol. This leaves little doubt decarboxylation would occur spontaneously in the malonate thiol ester - ammonium ion - thiol acetate complex of the condensing enzyme.

Discussion

The concerted decarboxylation-condensation model of the condensation reaction in polyketide biosynthesis was proposed by Arnstadt *et al* due to the lack of a primary deuterium isotope effect when dideutero-malonyl-CoA was used as the substrate for fatty acid synthetase, and the lack of tritium incorporation when the condensation was carried out in tritiated water.[9] These experiments were, however, of what might be termed "one-way" type[10a] in the sense that while one of the two possible outcomes would have been significant, the other is not. Thus if isotopic exchange had been observed, this would have provided strong evidence for participation by ionic intermediates. Failure to observe it could, however, be due to the factor noted above, i.e. that the special characteristic of enzyme reactions probably arise from the fact that absorption of the substrate into an active site involves displacement of solvent (water) from between them. Here, even if anions are involved, they clearly cannot undergo isotopic exchange with solvent if there is no solvent present.

As noted above, our calculations indicate that the decarboxylation and condensation cannot take place as a simple synchronous process because the calculated activation energy was much too large. This result could have been anticipated because the S_g2 reaction, being an autoactivated process due to the steric difficulties associated with forming a pentavalent adduct of

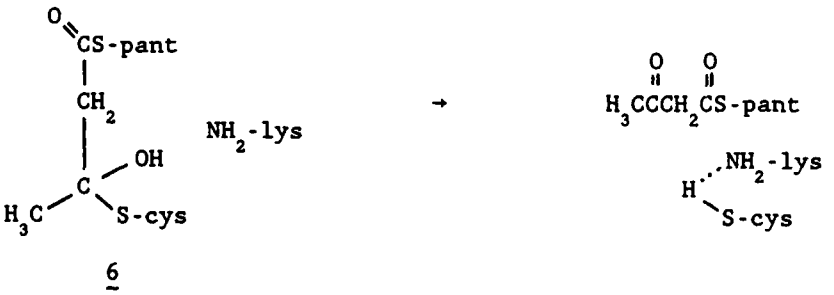
carbon,[47] cannot act as a driving force for formation of the reactants in it.

On the other hand, the calculations do imply that the condensation and decarboxylation are, at least, weakly concerted. The precise geometry of the malonate thiol ester, lysine residue and thiol acetate in the active site of the condensing enzyme is not known. It is possible the methene carbon of the malonate thiol ester is close enough to the carbonyl carbon of the thiol acetate at the time decarboxylation occurs that there is even greater coupling between the two reactions.

A model for the decarboxylation-condensation reaction of polyketide synthesis consistent with experimental evidence, then, involves the initial transfer of an acetyl group from acetyl-CoA to the active cysteine. The aqueous pKs of the thiol group of cysteine and the ϵ-NH_3^+ group of lysine are 8.18 and 10.53, respectively,[50] while the gas phase DPEs of alkyl thiols are on the order of 350 kcal/mol[32] and the PAs of alkyl amines are approximately 220 kcal/mol[37]. With the close proximity of the amino group perturbing the acidity of the thiol group, the lysine and cysteine residues will exist as the ammonium salt in the presence of water. The energy required to extrude the water during the preliminary stages of the acylation process, then, is provided by the exothermicity of the reverse proton transfer in the now solvent free environment. The association of the amino and thiol groups remains close,

however, and the proton is transferred back to the amino group during the actual acylation of the thiol group.

A malonyl group is then transferred from malonyl-CoA to the pant-SH. Whether the malonyl group is deprotonated after binding in the malonyl transferase or not until after transferring to the pant-SH is not known. However, the requirement for a solvent free environment for the decarboxylation dictates the malonyl-CoA be protonated when absorbed from solution.[10a,b] In the presence of the acylated cysteine residue and the protonated lysine residue, the deprotonated malonyl group decarboxylates and condenses with the acetyl carbonyl carbon. The resulting hemithiolacetal, 6,



6

then rearranges to the β-ketoacyl thiolester and the initial cys-SH···lys-NH_2 hydrogen bonded complex.

The inability to find a transition state for the proton relay condensation mechanism prevents a conclusive evaluation of this model, but does not eliminate this model as a possibility. The difficulty in finding the transition state for the proton relay mechanism could be the result of one or a combination of several

factors, both computational and chemical. Since several elementary processes are involved (i.e. formation of a carbon-carbon bond and two hydrogen transfers), an extensive grid search involving a separate coordinate for each process may be required to locate the transition state. Such a three dimensional grid search involving the model system in this study is not feasible.

The problem of locating a transition state may also have a chemical origin. As mentioned above, the actual geometries of the substrates in the enzyme are not known, and the possibilities are virtually endless. Of course, the location of a transition state will ultimately depend on selection of the proper geometries. While a large number of possibilities were tried, it is possible the "correct" one was not. Especially significant is the separation of the condensing carbons upon decarboxylation of the malonate thiol ester. If they are close enough, decarboxylation and condensation may be coupled enough to prevent any proton transfer. Additionally, although AM1 typically calculates activation barriers more accurately than MNDO, it is known to seriously overestimate barriers to some proton transfers.[12,49] In this case, that difficulty would be compounded by the reaction coordinate involving two proton transfers. Finally, at some point along the reaction coordinate a double-welled potential may develop, and proton tunnelling may play a significant role. In any of these cases, the transition state as visualized above may be extremely high in energy on the AM1 PE surface, and very difficult to locate.

Based on the results above, however, there is a simpler model which eliminates the above difficulties. As modelled, proton transfer does not occur until well after the transition state for decarboxylation is passed, and does not appear to play a significant role in the decarboxylation, aside from making the overall reaction more exothermic. While it is known the lys-NH_2 is close enough to the cys-SH and pant-SH to interact with them,[24] it is not known how close it is to the carbonyl oxygens of bonded acyl groups. As such, the effect of the ammonium group in the enzyme may be purely an electrostatic effect. The protonated amine group may be far enough away from the carbonyl group of the malonate that its positive charge initiates decarboxylation without actually transferring a proton in the process. Hence, the decarboxylation and condensation may take place without actual proton transfers except from the cys-SH to the lys-NH_2 and back. In this case, condensation should take place with only a small activation barrier, as with the enolate and acetate alone.

This further illustrates the difference between enzymatic reactions and reactions in aqueous solution. In the solvent free environment of the active sites of enzymes, any charge present will have a far-reaching effect on reactions. In solution, this effect would be severely attenuated by the dielectric properties of water.

Both of the above models are consistent with the inhibition of the condensing activity by iodoacetamide. Iodoacetamide is known to acylate the cys-SH:[50]

$$ICH_2CONH_2 + \text{cys-S}^{\text{lys-NH}_2\cdots\text{H}} \rightarrow \text{cys-SCH}_2CONH_2 + I^- + \text{lys-NH}_3^+$$

Whereas unmodified fatty acid synthetase in the absence of acetyl-CoA and NADPH decarboxylates malonyl residues very slowly, the enzyme inhibited by iodoacetamide decarboxylates malonyl-CoA much more rapidly, while at the same time preventing condensation.[50] Kresze *et al* attributed this effect to the confirmational change in the enzyme resulting from the binding of the iodoacetamide. While this may be part of the explanation, the binding of the iodoacetamide also results in the formation of the free ammonium ion required for decarboxylation. The carbonyl carbon of the iodoacetamide is less electrophilic than that of an acyl thiol ester, and the tetrahedral condensation product probably does not form to a significant extent. As a result, the enol or enolate eventually picks up a proton and is eliminated as acetyl-CoA, opening the way for the decarboxylation of another malonate residue.

References

1. Bu'lock, J. D. in *Comprehensive Organic Chemistry*, Haslam, E., Ed; Pergamon Press: Oxford, 1979; Vol. 5, pp. 927-946.

2. There is disagreement between authors as to whether fatty acids should be included in the category of polyketides (see, for example, ref. 1 and 3). For our purpose, we have followed the convention of ref. 1 and included them due to the similarity in biosynthetic pathways.

3. Herbert, R. B. *The Biosynthesis of Secondary Metabolites*, Chapman Hall: New York, 1981; Chapter 1.

4. Wood, H. C. S. in *Comprehensive Organic Chemistry*, Haslan, E., Ed; Pergamon Press: Oxford, 1979; Vol. 5, pp. 516-521.

5. Tsukamoto, Y.; Wong, H.; Mattick, J. S.; Wakil, S. J. *J. Biol. Chem.* 1983, *258*, 15312.

6. Suckling, C. J. *Chem. Soc. Rev.* 1984, *13*, 97.

7. Hall, C. M.; Wemple, J. *J. Org. Chem.* 1977, *42*, 2118, and references therein.

8. Bartmess, J. E.; Hays, R. L.; Caldwell, G. *J. Am. Chem. Soc.* 1981, *103*, 1338.

9. Arnstadt, K.-I; Schindlbeck, G.; Lynen, F. *Eur. J. Biochem.* **1975**, *55*, 561.

10. (a) Dewar, M. J. S., accepted for publication in *Enzyme*. (b) Dewar, M. J. S.; Storch, D. M. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 2225. (c) Dewar, M. J. S.; Storch, D. M. *J. Chem. Soc., Chem. Commun.* **1985**, 94.

11. Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899, 4907.

12. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

13. Dewar, M. J. S.; Reynolds, C. H. *J Comput. Chem.* **1986**, *7*, 140.

14. Available from Quantum Chemistry Program Exchange (QCPE).

15. (a) Fletcher, R.; Powell, M. J. D. *Comput. J.* **1963**, *6*, 163. (b) Davidon, W. C. *Ibid.* **1968**, *10*, 406.

16. Komornicki, A.; McIver, J. W., Jr. *J. Am. Chem. Soc.* **1972**, *94*, 2625.

17. A full force constant analysis requires the calculation of a complete Hessian matrix. Since this was not possible for complexes in which the freezing of selected internal coordinates was required to maintain separation of charged groups, the heats of formation

quoted for proposed minima are those from DFP optimized geometries. The heats of formation of proposed transition states, in these cases, are those of the approximate maxima located along the reaction coordinates in reaction path calculations and optimized by minimizing the gradient norm.[16,18]

18. (a) Bartels, R. H., University of Texas Center for Numerical Analysis, Report CNA-44, Austin, Texas, 1972. (b)Komornicki, A.; McIver, J. W. *Chem. Phys. Lett.* 1971, *10*, 303.

19. Dewar, M. J. S.; Ford, G. P. *J. Am. Chem. Soc.* 1977, *99*, 7822.

20. Pitzer, K. S.; Gwinn, W. D. *J. Chem. Phys.* 1942, *10*, 428.

21. (a) Shimakata, T.; Stumpf, P. K. *Arch. Biochem. Biophys.* 1982, *217* , 144. (b) Shimakata, T.; Stumpf, P. K. *Plant Physiol.* 1982, *69*, 1257.

22. Lynen, F. *Biochem. J.* 1967, *102*, 381.

23. (a) Singh, N.; Wakil, S. J.; Stoops, J. K. *Biochemistry* 1985, *24*, 6598. (b) Stoops, J. K.; Wakil, S. J. *J. Biol. Chem.* 1981, *256*, 8364. (c) Puri, R. N.; Porter, J. W. *Can. J. Biochem. Cell Biol.* 1985, *63*, 50. (d) Kumar, S. *J. Theor. Biol.* 1982, *95*, 263.

24. Wakil, S. J.; Stoops, J. K. in *The Enzymes*, 3rd Ed., Boyer, P.D., Ed.; Academic Press: New York, 1983; Vol. 16, pp 3-61.

25. Stoops, J. K.; Wakil, S. J. *Biochem. Biophys. Res. Commun.* 1982, *104*, 1018.

26. It should be noted that in this and the following schemes, hydrogen atoms are shown to indicate the ultimately required charge and mass balance. The actual dispositions of specific hydrogen atoms are not known, nor is the specific point in the synthesis at which deprotonation of the carboxyl group of the malonyl thiol ester occurs. This will be discussed later.

27. Simpson, T. J. *Biosynthesis* 1980, *6*, 1.

28. Stoops, J. K.; Henry, S. J.; Wakil, S. J. *J. Biol. Chem.* 1983, *258*, 12482.

29. Osterhelt, D.; Bauer, H.; Kresze, G.-B.; Steber, L.; Lynen, F. *Eur. J. Biochem.* 1977, *79*, 173.

30. See Chapter 3 of this dissertation.

31. The freezing of a limited number of internal coordinates is reasonable, considering the orientation of enzyme amino acid residues, and substrates bonded to the enzyme, is restricted.

32. Bartmess, J. E., intended for publication in *J. Phys. Chem. Ref. Data.*

33. Olivella, S.; Urpi, F.; Vilarrasa, J.*J. Comput. Chem.* 1984, *5*, 230.

34. Benson, S. W. *Thermochemical Kinetics*, 2nd Ed.; John Wiley & Sons: New York, 1976; Chapter 2.

35. Pedley, J. B.; Rylance, G. *Sussex-N.P.L. Computer Analysed Thermochemical Data: Organic and Organometallic Compounds*, Sussex University, 1977.

36. Stull, D. R.; Prophet, J. *JANAF Thermochemical Tables* NSRDS-NB337, 1971.

37. Lias, S. G.; Liebman, J. F.; Levin, R. D. *J. Phys. Chem. Ref. Data* 1984, *13*, 695.

38. Dewar, M. J. S.; Rzepa, H. S. *J. Am. Chem. Soc.* 1978, *100*, 784.

39. Cox, J. D.; Pilcher, G. *Thermochemistry of Organic and Organometallic Compounds*, Academic Press: London, 1970.

40. Due to the large error in the calculated $\Delta H_f(H^+)$ (331.5[33] kcal/mol MNDO, 314.9[30] kcal/mol AM1, 367.2[36] kcal/mol exptl), the experimental value is used in determining calculated values of PAs and DPEs.

41. Jencks, W. P. *Catalysis in Chemistry and Enzymology*, McGraw Book Company: New York, 1969; pp. 116-120.

42. Storch, D. M., Ph.D. Dissertation, The University of Texas at Austin, 1985.

43. The calculated barrier to proton transfer may not be real. Since the proton transfer did not result in decarboxylation, there was no accumulation of charge on the carbonyl oxygen, as would occur if decarboxylation was occurring simultaneously with proton transfer. Furthermore, proton transfer from methyl ammonium ion to acetone is predicted by AM1 to occur endothermically, as would be expected given the relative basicities of acetone and methyl amine,[30,37] but without passing through a transition state.[44] This is consistent with *ab initio* predictions of negligible barriers to proton transfer between oxygen and nitrogen bases aside from the endothermicity of the transfer itself.[45] Indeed, based on these results even the barrier to proton transfer in the malonate alcohol ester - ammonium ion complex may be partly the result of the restricted geometry.

44. Dieter, K. M. Unpublished work.

45. Cao, H. Z.; Allavena, M.; Tapia, O.; Evleth, E. H. *J. Phys. Chem.* **1985**, *89*, 1581, and references therein.

46. Blackburn, S. *Enzyme Structure and Function*, Dekker: New York, 1976, pp. 62-7.

47. Dewar, M. J. S.; Healy, E. F. *Organometallics* 1982, *1*, 1705.

48. White, A.; Handler, P.; Smith, E. L. *Principles of Biochemistry*, 5th Ed.; McGraw-Hill Book Company: New York, 1973, p. 104.

49. For this same reason, the direct proton transfer from the enol to the acyl carbonyl group cannot be conclusively eliminated.

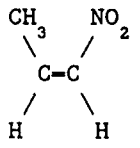
50. Kresze, G. -B.; Steber, L.; Oesterhelt, D.; Lynen, F. *Eur. J. Biochem.* 1977, *79*, 191.

Chapter 5

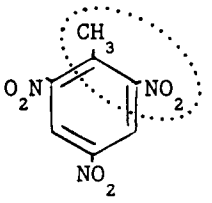
AN AM1 STUDY OF THE MASS SPECTRAL FRAGMENTATION OF CIS-1-NITROPROPENE

Introduction

In a recent theoretical study,[1] Turner and Davis proposed cis-1-nitropropene (C-1-NP), 1, as a suitable model for studying possible mechanisms of the bulk phase thermolysis of 2,4,6-trinitrotoluene (TNT), 2.



1



2

Their choice was based on the similarity between the structure of C-1-NP and the structure of the reactive center of TNT (dotted portion of 2) in thermolysis. Using MINDO/3[2] and MNDO,[3] they looked at intra- and intermolecular hydrogen transfer from the methyl group to the nitro group, intra- and intermolecular oxygen insertion from the nitro group into a C-H bond of the methyl group, rearrangement of the nitro group to a nitrite group, and dissociation of NO_2 by simple R-NO_2 bond scission. The energetic

ordering of the calculated heats of reaction were nearly identical when C-1-NP and TNT were the reactants, providing hope that the energetic ordering of the transition states of the model system would also parallel those of TNT.[1] While experimental evidence for TNT was most consistent with intermolecular hydrogen transfer,[1] the calculated activation energies for all C-1-NP reactions except NO_2 dissociation were reasonably close to the observed activation energy for TNT thermolysis. The calculated R-NO_2 bond dissociation enthalpy (BDE), 65 kcal/mol,[4] was 15-30 kcal/mol larger than the MINDO/3 calculated activation enthalpies for the other intramolecular processes.

Dewar *et al* similarly found simple R-NO_2 bond scission to be unfavorable in the thermolysis of nitromethane.[5] Again using MINDO/3, they showed that a nitro to nitrite rearrangement, followed by dissociation of NO, was more consistent with gas phase kinetic data than was the direct dissociation of NO_2.

These findings are in direct contrast to gas phase pyrolysis results for substituted nitrobenzenes.[6] Under conditions in which surface catalyzed reactions were minimized, Gonzalez *et al* determined the major process involved scission of the Ar-NO_2 bond, although the nitro to nitrite rearrangement and subsequent NO dissociation also occurred to a lesser extent. This was true for all compounds, even those with a methyl group ortho to the departing nitro group, and despite the fact the R-NO_2 BDE derived by Gonzalez

(70 kcal/mol) was even larger than that calculated by Turner and Davis.[1] Indeed, the largest Arrhenius preexponential factor and smallest Arrhenius activation energy derived from the nitrobenzene results were for an ortho methyl derivative, and this effect was attributed to the steric effects of the methyl group. At the same time, there was no evidence of direct participation of the methyl group in the decomposition.

The dissociation of NO_2 and the nitro to nitrite rearrangement with subsequent NO dissociation are known to also play a role in the decomposition of molecular ions of nitro compounds. Both decomposition channels, for example, have recently been shown to be active in the fragmentation of nitromethane[7] and nitrobenzene[8] molecular ions.

In the mass spectral fragmentation of the C-1-NP molecular ion ($C\text{-}1\text{-}NP^{+\cdot}$), these reactions would compete with fragmentation involving direct participation of the methyl group. Given the above theoretical and experimental results for neutral nitro compounds, it is of interest to determine if methyl group participation in the fragmentation of $C\text{-}1\text{-}NP^{+\cdot}$ does occur, and to compare the calculated enthalpic requirements of the competing processes.

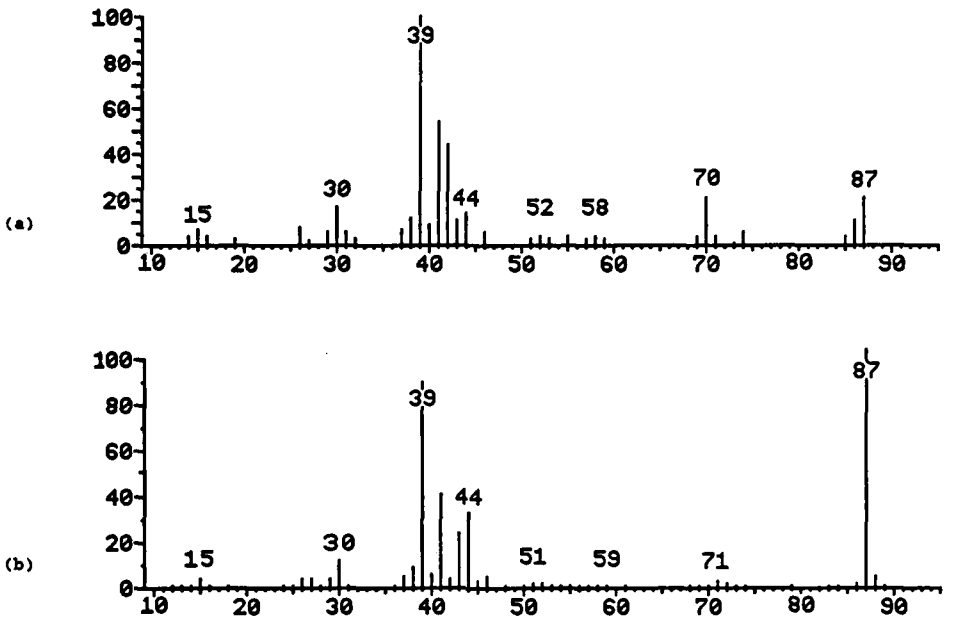


Figure 1. The 70 eV electron impact mass spectra of (a) cis-1-nitropropene and (b) trans-1-nitropropene.[9] The vertical axis is relative abundance and the horizontal axis is mass-to-charge ratio.

Mass Spectra of Cis- and Trans-1-Nitropropene

The 70 eV electron impact (EI) mass spectra[9] of C-1-NP and trans-1-nitropropene(T-1-NP) are shown in Figure 1, and Table 1 lists the relative abundances of the peaks. There are notable similarities in the two spectra. In particular, the molecular ion contribution to total ion abundance is relatively large, m/z 87 being the base peak for T-1-NP while having a relative abundance of 21.4 for C-1-NP. The series of peaks from m/z 37 to m/z 44 are also very similar with the exception of the peak at m/z 42.

Table 1. Relative abundances in the 70 eV EI mass spectra of cis- and trans-1-nitropropene.[a,b]

m/z[c]	cis	trans	m/z[c]	cis	trans
15	6.7	3.6	41	53.9	40.9
26	7.8	3.9	42	43.5	4.3
29	6.4	3.9	43	11.0	24.7
30	17.1	12.2	44	13.6	33.1
31	6.4	0.2	46	5.5	4.6
37	7.2	5.1	70	21.2	0.7
38	11.6	9.4	74	6.4	1.2
39	100.0	89.7	86	11.3	1.9
40	9.0	6.2	87	21.4	100.0

[a]Spectra from ref.9 [b]Peaks are listed only if the relative abundance for at least one of the two isomers exceeds 5.0. [c]Mass to charge ratio.

Additionally, there is significant fragmentation of $C_3H_5O + NO^+$ (m/z 30) and $C_3H_5^+ + NO_2$ (m/z 41) in both isomers. Loss of NO and NO_2^+ (m/z 57 and m/z 46, respectively) appears to be relatively minor, assuming negligible secondary fragmentation.

The spectra clearly have differences, however. This indicates cis-trans isomerization does not occur to a significant extent, a somewhat surprising result considering AM1 predicts the barrier to rotation around the C^1-C^2 bond to be only 11 kcal/mol.[10] The spectra also show more extensive fragmentation of C-1-NP$^{+\cdot}$ than the T-1-NP molecular ion (T-1-NP$^{+\cdot}$). This is evident from the difference in the relative abundances of the molecular ions, and from the larger number of significant peaks (relative abundance >5.0) in the C-1-NP spectrum. Of particular note is the presence of

an m/z 70 peak in the C-1-NP spectrum, presumably from the loss of a hydroxyl radical from C-1-NP$^{+\cdot}$, and its absence from the T-1-NP spectrum. The implication of these results is that there is direct methyl group participation in the fragmentation of an m/z 87 ion. As will be seen, the m/z 86 peak is also significant in this regard.

Calculation Procedure

The MNDO, MINDO/3, and AM1[11] semiempirical quantum mechanical models have been shown to give results comparable to *ab initio* models using split basis sets in a fraction of time required for the *ab initio* calculations.[12] While MNDO is known to have difficulties with compounds containing nitro groups,[1,3,5] the MINDO/3 average absolute errors for these compounds is essentially the same as for all CHON compounds (approximately 9 kcal/mol).[1] Consequently, prior to the development of AM1, MINDO/3 was the method of choice in studies of nitro-containing compounds.[1,5] However, MNDO and MINDO/3 both predict the nitro group of C-1-NP and the C-1-NP$^{+\cdot}$ to be perpendicular to the plane of the alkene. AM1 predicts the nitro group of both species to be in the plane, as one would intuitively expect. Additionally, AM1 results for nitro compounds indicate an accuracy of the same order as for MINDO/3.[1,11,13] Since conjugation of the nitro group with the alkene π bond, as well as the positions of the oxygen atoms in the plane of the alkene as opposed to out of the plane, could possibly have a significant effect on calculated reaction profiles, AM1 was used for

this study. Due to the unusually large errors in the AM1 heats of formation of NO, NO_2, and the larger than average errors for NO^+, NO_2^+, and OH, experimental heats of formation for these species are used in calculating heats of reactions for reactions involving them.

All calculations were carried out using AM1 as incorporated in the AMPAC package of computer programs[14] modified to include additional gradient norm minimization routines. All geometries were optimized by minimizing the energy with respect to all geometrical variables using the DFP method.[15] Approximate transition states located by reaction path calculations were optimized by minimizing the gradient norm.[16-18] True minima and transition states were characterized by the presence of *exactly* zero or one negative eigenvalue, respectively, in corresponding Hessian matrices.[16a]

All calculations involving radicals were initially attempted in the restricted Hartree-Fock (RHF) formalism with the half-electron (HE) approximation.[19] Reaction path calculations for many of the reactions studied gave discontinuous results using this method. The spin-unrestricted version of AM1 was also tried, but did not eliminate the discontinuities. As a result, 2x2 configuration interaction (CI) (involving the highest doubly occupied MO (HDMO) and the singly occupied MO (SOMO)) or 3x3 CI (involving the HDMO, SOMO and lowest unoccupied MO (LUMO)) was required to obtain continuous reaction path results. In all cases but one, the CI served only to reach regions of the geometric

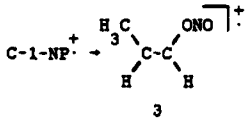
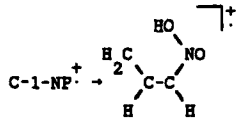
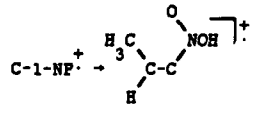
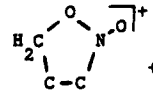
surface that were unattainable using single determinantal wave functions. In these cases, the approximate transition states eventually located showed no mixing of configurations, and were optimized satisfactorily with the RHF/HE method. The exception was the dissociation of a hydroxyl radical from the product of hydrogen transfer to the nitro group, for which 3x3 CI was required to locate and optimize the transition state. Results are given in Table 2.

Results

A. Ionization of C-1-NP. The heat of formation of the product of vertical ionization of C-1-NP was calculated as 256.8 kcal/mol, while that of the optimized C-1-$NP^{+\cdot}$ was 248.8 kcal/mol. The only significant geometrical change in the relaxation of C-1-$NP^{+\cdot}$ was the equalization of the two carbon-carbon bonds (1.34 Å and 1.47 Å in C-1-NP, 1.43 Å and 1.42 Å in C-1-$NP^{+\cdot}$). This was consistent with a singly occupied carbon-carbon π orbital in the cation, as the eigenvectors indicated.

B. Dissociation of NO_2 and NO_2^+ from C-1-$NP^{+\cdot}$. Initial attempts to locate a transition state (TS) for the dissociation of NO_2 or NO_2^+ using RHF/HE calculations and treating the R-NO_2 bond as the reaction coordinate resulted in discontinuous results. Analysis of the vectors showed a reversal of the HDMO and SOMO at the point of discontinuity. Using 2x2 CI, an approximate TS was located which had no mixing of configurations. The TS then optimized with the

Table 2. AM1 calculated heats of formation (ΔH_f), heats of activation ($\Delta H^{\neq}$) and heats of reaction (ΔH_{rxn}) for the mass spectral fragmentation of cis-1-nitropropene (kcal/mol).

Reaction	ΔH_f Reactant	TS	Product	$\Delta H^{\neq}$	ΔH_{rxn}
C-1-NP → C-1-NP$^{+\cdot}$	5.9		248.8[a]		242.9
C-1-NP$^{+\cdot}$ → $C_3H_5^+$ + NO_2	248.8[a]	266.2	259.1[b]	17.4	10.3[b]
→ C_3H_5 + NO_2^+			287.6[b]		38.8[b]
C-1-NP$^{+\cdot}$ → 	248.8[a]	268.0	189.8	19.2	-59.8
3 → C_3H_5O + NO^+	189.8	-	222.0[b]	-	32.2[b]
→ $C_3H_5O^+$ + NO		-	212.2[b]	-	22.4[b]
C-1-NP$^{+\cdot}$ → 	248.8[a]	259.0	229.9	10.2	-18.9
4 → $C_3H_4NO^+$ + OH	229.9	255.3[c]	244.9[b]	25.4[c]	15.0[b]
C-1-NP$^{+\cdot}$ → 	248.8[a]	298.6	246.5	49.8	-2.3
C-1-NP$^{+\cdot}$ →  +H	248.8[a]	286.5	272.9	37.7	24.1

[a] ΔH_f for the optimized C-1-NP$^{+\cdot}$. [b] The experimental ΔH_f^{20} values for NO_2 (7.9 kcal/mol), NO_2^+ (231.3 kcal/mol), NO (21.6 kcal/mol), NO^+ (236.6 kcal/mol) and OH (9.3 kcal/mol) are used in these calculations. [c] Ref. 21.

RHF/HE method at 1.8 Å, with a C-C-C-N dihedral angle of 12° (as opposed to 0° in C-1-NP$^{+\cdot}$). The CI wavefunction evidently enabled the breaking of the planar symmetry and, as a result, a smooth transition of orbitals from reactants to products. The lone negative frequency in the force constant analysis was -823 cm^{-1}, and the components of the corresponding normal vector primarily involved motion of C^1 and the hydrogens bonded to C^1 and C^2. This motion would ultimately restore planarity to the hydrocarbon fragment and increase the bond angle of the hydrogen bonded to C^1 from 123° to near 180° in the $C_3H_5^+$ product. The enthalpy of activation and enthalpy of reaction for the formation of $C_3H_5^+ + NO_2$ were 17.4 and 10.3 kcal/mol, respectively. Dissociation to produce C_3H_5 and NO_2^+ is predicted to be 30 kcal/mol more endothermic.

C. **Nitro to nitrite rearrangement.** The approximate TS was located with RHF/HE calculations by treating the C^1-N-O bond angle for the oxygen closest to the methyl group as the reaction coordinate. In the optimized TS, Figure 2, the nitro group is virtually perpendicular to the plane of the alkene. The incipient C-O bond is 1.7 Å, while the corresponding N-O bond has stretched only from 1.20 Å in C-1-NP$^{+\cdot}$ to 1.26 Å. In the product, 3, the C-O bond is 1.30 Å and the O-NO bond is extremely long, 1.69 Å. The negative frequency of the force matrix is -705 cm^{-1}, the components of the vector primarily involving motion of C^1 and the NO_2 group, although there is also significant participation of the hydrogen bonded to C^2. The rearrangement is predicted to be very exothermic

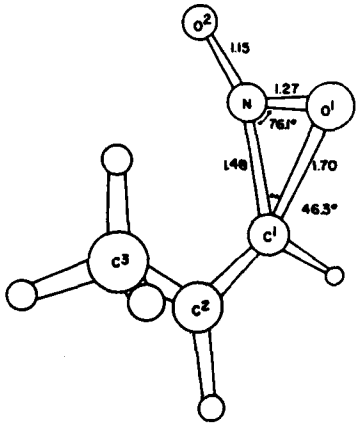


Figure 2. Transition state for the nitro to nitrite rearrangement of C-1-NP^{+}. The O^{1}-N-C^{1}-C^{2} dihedral angle is 94.4°. In this and the following figures, bond lengths are in angstroms and bond angles are in degrees. Atoms not labeled are hydrogens.

(ΔH_{rxn} = -59.8 kcal/mol) and the enthalpy of activation, 19.2 kcal/mol, is similar to that for NO_2 dissociation. The strain associated with the three membered cyclic TS for the rearrangement should decrease the probability of the rearrangement occurring relative to loss of NO_2. When it does occur, however, the 80 kcal/mol of reverse activation energy stored in the product makes rapid decomposition of 3 likely.

D. Dissociation of NO and NO^{+} from 3. No TS could be located for dissociation of NO or NO^{+} from 3. This is reasonable considering there is very little geometry change expected in the fragments due to the dissociation. The 22 and 32 kcal/mol required to dissociate NO and NO^{+}, respectively, are certainly insufficient to prevent the fragmentations given the vibrational energy stored in 3 as a result of the nitro to nitrite rearrangement. At first

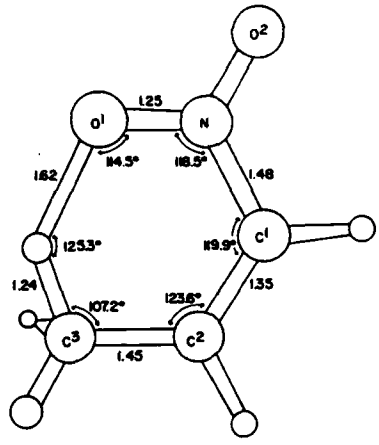


Figure 3. Transition state for hydrogen transfer from C^3 of C-1-$NP^{+\cdot}$ to O^1. The $H\text{-}C^3\text{-}C^2\text{-}C^1$ dihedral angle is -30.4°, and the $O^1\text{-}N\text{-}C^1\text{-}C^2$ dihedral angle is 8.6°.

glance it seems surprising the fragmentation to $C_3H_5O^+$ + NO requires less energy than the fragmentation to C_3H_5O + NO^+, since there is only a very small m/z 57 peak (relative abundance 3.2) while there is a sizeable m/z 30 peak (relative abundance 17.1). This will be discussed further in the Discussion Section.

E. Hydrogen transfer from the methyl group to the nitro group. As with the NO_2 dissociation from C-1-$NP^{+\cdot}$, 2x2 CI calculations were required to eliminate discontinuities observed for RHF/HE calculations treating the H-O distance as the reaction coordinate. The approximate TS located was then optimized with the RHF/HE method, as there was no mixing of configurations at the predicted TS. The H-O distance in the optimized TS, Figure 3, is 1.62 Å, while the H-C bond length increased from 1.14 to 1.24 Å and the N-O bond length from 1.20 to 1.25 Å. The oxygen atom and

hydrogen atom were 0.5 and 0.6 Å out of the C-C-C plane, respectively. The negative frequency of the force matrix was -842 cm^{-1}, the motion of the normal vector being composed of motion of the transferring hydrogen, as well as adjustment of the other hydrogens bonded to C^3. The 10.2 kcal/mol activation barrier is lower than those for NO_2 dissociation and nitro to nitrite rearrangement. The formation of the TS does limit the motion of the methyl rotor, resulting in an unfavorable entropic contribution to the formation of the TS. At least 30 kcal/mol of reverse activation energy is stored in the product, 4, as a result of the hydrogen transfer.

F. Dissociation of OH from 4. A 3x3 CI calculation was required to locate and optimize the TS for the dissociation of the OH radical from 4. In the TS, the N-OH bond length was 2.0Å and the O-N-C-C dihedral was 6°, as compared to 1.3 Å and 1°, respectively, in 4. The negative frequency in the force matrix was -443 cm^{-1}, and the normal vector was composed primarily of the motion of the OH group. The 25.4 kcal/mol activation barrier[21] was more than compensated for by the reverse activation energy from the rearrangement, which was stored as vibrational energy in 4.

G. Hydrogen transfer from C^1 to the nitro group. While the presence of the m/z 70 peak in the C-1-NP spectrum is interpreted as a sign of interaction between the methyl and nitro groups during fragmentation, these same peaks could result from hydrogen transfer

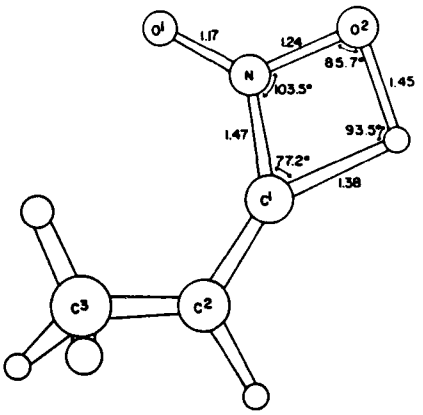


Figure 4. Transition state for hydrogen transfer from C^1 of C-1-$NP^{+\cdot}$ to O^2. The $H\text{-}C^1\text{-}C^2\text{-}C^3$ dihedral angle is 132.7°, while the $N\text{-}C^1\text{-}C^2\text{-}C^3$ dihedral angle is -40.3°.

from C^1 to the nitro group. This transfer could occur in both C-1-NP or T-1-NP. If the enthalpic requirement for this reaction was similar to that for hydrogen transfer from the methyl group, a different explanation might be required for the differences in the two spectra. As with the previous hydrogen transfer, 2x2 CI was required to locate the TS, but it was optimized satisfactory with the RHF/HE method. The TS, Figure 4, was extremely strained. The negative frequency was -2244 cm^{-1} and corresponded entirely to the motion of the migrating hydrogen. The high activation barrier, 49.8 kcal/mol, plus the strain in the transition state make this rearrangement highly improbable, supporting the previous interpretation of the m/z 70 peak.

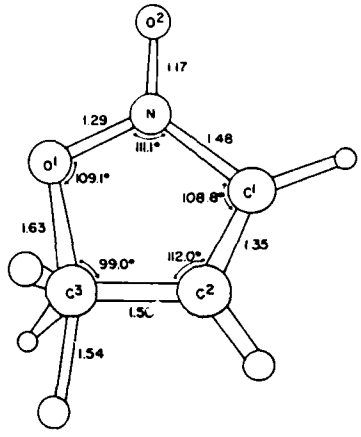


Figure 5. Transition state for the loss of a hydrogen atom from C^1 of C-1-$NP^{+\cdot}$. The O^1-C^3-C^2-C^1 dihedral angle is 0.0°.

H. Dissociation of a hydrogen atom from the methyl carbon. It is interesting to note the significant m/z 86 peak (relative abundance 11.3) in the C-1-NP spectrum and the negligible peak (relative abundance 1.9) in the T-1-NP spectrum. When the C^3-O distance was treated as the reaction coordinate and incrementally decreased, a hydrogen atom eventually dissociated from the methyl group and a cationic nitrogen analog to a lactone was formed. The same product resulted when the length of a C-1-$NP^{+\cdot}$ H-C^3 bond was incrementally increased. The TS for this reaction was located by starting with the product of the dissociation and incrementally decreasing the H···C^3 distance. The RHF/HE method was used to locate and optimize the TS, Figure 5. The H-C^3 distance in the TS was 1.54 Å, while the incipient C^3-O bond was 1.63 Å. The 37.7 kcal/mol activation is significantly larger than those of the other

occurring reactions, and formation of this TS again restricts the motion of the methyl rotor.

Discussion

Electron impact ionization produces molecular ions with internal energies ranging from zero to over 20 eV, the average being a few eV.[22] According to the quasi-equilibrium theory,[23] the statistical theory generally accepted as describing the fragmentation in mass spectrometers,[22,24] this internal energy is rapidly and randomly distributed among the internal degrees of freedom in the ground state of the molecular ion. Fragmentation is then statistical, specific fragmentations occurring only when the nuclei are in the proper configuration and sufficient vibrational energy is concentrated in the necessary degrees of freedom.[23] Arguments involving fragmentation from isolated electronic states and nonrandom distribution of vibrational energy are sometimes used to explain apparently nonstatistical fragmentation.[8,25] Nitromethane and nitrobenzene are two compounds for which these arguments have been proposed.[8] It is difficult to eliminate alternative explanations such as noninterconverting structural isomers,[8,26] however, and incomplete randomization has not been proven directly.[8] Consequently, it is assumed C-1-NP$^{+\cdot}$ fragments statistically from its ground electronic state. Additionally, while there is experimental evidence that some fragmentations formally described as direct bond cleavages actually involve more extensive intramolecular

interactions,[24] including dissociation of NO_2 from nitrobenzene,[8] all dissociations in this study are assumed to involve loose transition states.

Attempts to rationalize low m/z peaks are often futile because they are likely to correspond to products of secondary fragmentation, and are also more likely to involve extensive rearrangements. The exception is when there is an obvious, straight-forward path to the low m/z fragment, such as m/z 30 (C_3H_5O + NO^+) and m/z 41 ($C_3H_5^+$ + NO_2), and even then there is no guarantee this path accounts for the full peak abundance. Higher m/z fragments, while still possibly involving rearrangement, will more probably correspond to products of fragmentation of the molecular ion or its isomer. The m/z 86 and m/z 70 peaks, which are present in the C-1-NP spectrum and virtually absent in the T-1-NP spectrum, fall in this latter category, and are easily rationalized from AM1 results on the basis of direct participation of the methyl group in the fragmentation of C-1-$NP^{+\cdot}$. The AM1 results also explain certain similarities in the two spectra.

In both spectra, there are large m/z 41 peaks corresponding to fragmentation to $C_3H_5^+$ + NO_2, and small m/z 46 peaks, corresponding to fragmentation to C_3H_5 + NO_2^+. This would certainly be expected on the basis of AM1 calculations, which indicate it is more favorable for the hydrocarbon fragment to carry the charge by 30 kcal/mol. While the formation of $C_6H_5^+$ + NO_2 from the

nitrobenzene molecular ion apparently occurs with no reverse activation energy,[8] there is a small reverse activation energy for the C-1-NP$^{+\cdot}$ fragmentation. This is reasonable since there is significant geometric relaxation of the $C_3H_5^+$ fragment due to fragmentation, whereas small changes would be expected in the relatively rigid phenyl cation.

On the basis of the calculated enthalpy of activation for the nitro to nitrite rearrangement, significant rearrangement should occur, and, as indicated by the size of the m/z 30 peak, does. The negative frequency of the TS is relatively low, indicating the strain in the three membered cyclic TS is not excessively large and is insufficient to make the rearrangement uncompetitive.

The AM1 results for fragmentation of 3 to $C_3H_5O^+$ + NO and C_3H_5O + NO^+ do seemingly conflict with the presence of m/z 30 peaks and the absence of m/z 57 peaks in the C-1-NP and T-1-NP spectra. These results are, however, perfectly consistent. The nitro to nitrite rearrangment occurs with a large reverse activation energy, resulting in approximately 80 kcal/mol of vibrational energy being stored in 3. Upon fragmention of 3, the excess energy is partitioned between the two fragments, with the larger fragment expected to carry off more of the energy (degree of freedom effect[23]). As a result, the $C_3H_5O^+$ ion will likely fragment further. One possibility is that it loses a water molecule, thereby contributing to the large m/z 39 peaks in both spectra. On the

other hand, fragmentation of NO^+ is extremely unlikely and most NO^+ formed is probably detected.

The hydrogen transfer from the methyl group to the nitro group has the lowest activation barrier of all the reactions studied, and results in at least 30 kcal/mol of vibrational energy being stored in 4. This is more than enough energy to result in subsequent decomposition to 5 + OH,[21] and the unfavorable entropy factor associated with the restriction of the methyl rotor in the TS does not outweigh the relatively favorable enthalpic factor. These results, plus the high energy and tight TS required for hydrogen transfer from C^1 to the nitro group, clearly explain the presence of the m/z 70 peak in the C-1-NP spectrum and its absence in the T-1-NP spectrum.

Given the energetic requirements for the hydrogen transfer and hydroxyl radical dissociation versus those for the dissociation of NO_2, it is even a little surprising the m/z 70 peak is less than half the size of the m/z 41 peak in the C-1-NP spectrum. One reason may be the restriction of the methyl rotor in the TS for hydrogen transfer. Additionally, there is another route to the production of an m/z 41 peak, namely loss of NO_2 from 3. The heat of formation for the TS of this process would be expected to be similar to that for the loss of NO_2 from C-1-$NP^{+\cdot}$. This is probably a minor process, however, since loss of NO and NO^+ is much more heavily favored energetically, as would be anticipated from the long O-NO

bond. In any case, it is unlikely these possibilities provide the full explanation.

On the other hand, a major difference between the C-1-NP and T-1-NP spectra still unaccounted for is the m/z 42 peak. Its large size in the C-1-NP spectrum and virtual absence in the T-1-NP spectrum indicates its generation in some way involves interaction between the methyl and nitro groups. Only three combinations of atoms present in C-1-NP add up to an m/z 42: $C_2H_2O^+$, $C_2H_4N^+$ and CNO^+. A reasonable source of the m/z 42 ion consistent with the above data, then, would be fragmentation of 5 to ethylene and CNO^+. While the total heat of formation for these products is predicted to be quite high (CNO^+ 342.4 kcal/mol (AM1), C_2H_4 12.5 kcal/mol (exptl[20]), it is still reachable, considering the internal energy deposited in the molecular ion during ionization. Other explanations for the m/z 42 peak are either inconsistent with the contrast between the C-1-NP and T-1-NP spectra or involve significantly more rearrangement.

Finally, the dissociation of a hydrogen atom from the methyl group with simultaneous formation of the five membered cyclic ion is also consistent with the two spectra, i.e the much larger m/z 86 peak in the C-1-NP spectrum as opposed to the T-1-NP spectrum. While the enthalpy of activation and endothermicity for this reaction is significantly larger than the other occurring intramolecular processes, they are less than would be expected if a

simple C-H bond scission was occurring without the assistance of other intramolecular interactions.

Conclusions

Differences in the mass spectra of C-1-NP and T-1-NP imply there is significant involvement of the methyl group in the fragmentation of C-1-NP$^{+\cdot}$. The results of AM1 calculations for several possible fragmentation paths support this conclusion and appear to provide a good rationalization for the major features of the C-1-NP spectrum. Further experimental work, involving softer ionization and collision induced dissociation, is necessary to confirm the proposed looseness of the transition states for dissociations and secondary fragmentation pathways.

Based on these results, it can be concluded that the MINDO/3 and MNDO results of Turner and Davis[1] for possible thermal decomposition mechanisms for C-1-NP are probably also correct. As noted by them, however, using C-1-NP as a model for TNT has a shortcoming in that the effects on the aromaticity of TNT is not represented in calculations involving C-1-NP. This is the most likely reason the R-NO_2 bond scission, predicted to be unfavorable for C-1-NP, was found to be the dominant process in the gas phase pyrolysis of substituted nitrobenzenes.

References

1. Turner, A. G.; Davis, L.P. *J. Am. Chem. Soc.* 1984, *106*, 5447.

2. Bingham, R. C.; Dewar, M. J. S.; Lo, D. H., *J. Am. Chem. Soc.* 1975, *97*, 1285, 1294, 1302, 1307.

3. Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* 1977, *99*, 4899, 4907.

4. Due to the MNDO underestimation of stability of nitroaromatics and the MINDO/3 overestimation of stability of NO_2, the actual MNDO and MINDO/3 calculated R-NO_2 BDEs (29.8 and 41.4 kcal/mol, respectively) were judged to be too small. This value, then, is the MINDO/3 value corrected for the known error for NO_2.[1]

5. Dewar, M. J. S.; Ritchie, J. P.; Alster, J. *J. Org. Chem.* 1985, *50*, 1031.

6. Gonzalez, A. C.; Larson, C. W.; McMillen. D. F.; Golden, D. M. *J. Phys. Chem.*, 1985, *89*, 4809.

7. Gilman, J. P.; Hsieh, T.; Meisels, G. G. *J. Chem. Phys.* 1983, *78*, 1174.

8. Nishimura, T.; Das, P. R.; Meisels, G. C. *J. Chem. Phys.* 1986, *84*, 6190.

9. Turner, A. G.; DeLong, M. J.; Pflug, J. L., Frank J. Seiler Research Laboratory, USAF Academy, CO., unpublished results. Identification of the isomers was accomplished using NMR.

10. The estimated rotation barrier is based on the ΔH_f of optimized C-1-$NP^{+\cdot}$ (C-C-C-N dihedral angle 0°) and the ΔH_f for the cation with the C-C-C-N dihedral angle fixed at 90°.

11. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

12. (a) Dewar, M. J. S., *J. Phys. Chem.*, **1985**, *89*, 2145. (b) Dewar, M. J. S.; Storch, D. M. *J. Am. Chem. Soc.* **1985** *107*, 3898.

13. See Chapter 3 of this dissertation.

14. Available from Quantum Chemistry Program Exchange (QCPE).

15. (a) Fletcher, R.; Powell, M. J. D. *Comput. J.* **1963**, *6*, 163. (b)Davidon, W. C. *Ibid.* **1968**, *10*, 406.

16. (a) Komornicki, A.; McIver, J. W., Jr. *J. Am. Chem. Soc.* **1972** *94*, 2625. (b) Komornicki, A.; McIver, J. W. *Chem. Phys. Lett.* **1971**, *10*, 303.

17. Bartels, R. H., University of Texas Center for Numerical Analysis, Report CNA-44, Austin, Texas, 1972.

18. Powell, M. J. D. in *Numerical Methods for Nonlinear Algebraic Equations* Rabinowitz, P., Ed.; Gordon & Breach Science Publishers: New York, 1970; Chapters 6 and 7.

19. Dewar, M. J. S.; Hashmall, J. A.; Venier, C. G. *J. Am. Chem.Soc.* **1968**, *90*, 1953. (b) Dewar, M. J. S.; Trinajstic, N. *J. Chem. Soc. D:* **1970**, 646. (c) Dewar, M. J. S.; Trinajstic, J. *J. Chem. Soc. A:* **1971**, 1220.

20. Wagman, D. D.; Evans, W. H.; Parker, V. B.; Schumm, R. H.; Halow, I.; Bailey, S. M.; Churney, K. L.; Juttall, R. L. The NBS Tables of Chemical Thermodynamic Properties: Selected Values for Inorganic and C_1 and C_2 Organic Substances in S.I.Units. *J. Phys. Chem. Ref. Data* **1982**, *11*, Suppl. 2.

21. The barrier height for OH dissociation from $\underset{\sim}{4}$ is probably underestimated due to overcompensation for electron correlation in CI calculations. The corrected heat of formation for the TS would be expected to be significantly less than the RHF/HE heat of formation for the TS optimized with CI, which is 286.0 kcal/mol. Therefore, this process should be very competitive with the other processes.

22. McLafferty, F. W. *Interpretation of Mass Spectra, 3rd Ed.*, University Science Books: Mill Valley, CA, 1980; Chapter 7.

23. Rosenstock, H. M.; Wallenstein, M. B.; Wahrhaftig, A. L.; Eyring, H. *Proc. Natl. Acad. Sci. USA* **1952**, *38*, 667.

24. Levsen, K. *Fundamental Aspects of Organic Mass Spectrometry*, Verlag Chemie: Weinheim, 1978; Chapters 2 and 3.

25. Lifshitz, C. *J. Phys. Chem.* **1983**, *87*, 2304.

26. For an example of opposing arguments for the case of methyl nitrite, see: (a) Baer, T.; Hass, J. R. *J. Phys. Chem.*, **1986**, *90*, 451. (b) McKee, M. L. *J. Phys. Chem.* , **1986**, *90*, 2335.

APPENDIX 1

ARCHIVE OF SELECTED GEOMETRIES FROM CHAPTER 4

METHYL MALONATE, HYDROGEN BOND ON CARBONYL FROM CH3NH3+
REACTANT FOR DECARBOXYLATION, -134.3 KCAL/MOL AM1

O	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.417235	1	0.000000	0	0.000000	0	1	0	0
H	1.117674	1	110.735084	1	0.000000	0	2	1	0
H	1.118597	1	104.300045	1	119.497198	1	2	1	3
H	1.118561	1	110.019539	1	-121.834838	1	2	1	3
C	1.371443	1	118.026232	1	55.300773	1	1	2	3
O	1.244877	1	117.075765	1	1.075932	1	6	1	2
H	1.852926	1	132.951096	0	25.241995	0	7	6	1
N	1.048692	1	131.415604	0	168.219380	0	8	7	6
H	1.021932	1	107.296805	1	52.413308	1	9	8	7
H	1.028082	1	105.766459	1	-62.976110	1	9	8	7
C	3.940243	1	137.580158	1	24.274125	0	7	6	1
H	1.122720	1	109.960066	1	178.154915	1	12	9	8
H	1.122852	1	109.516381	1	-61.527881	1	12	9	8
H	1.124220	1	109.266493	1	58.163676	1	12	9	8
C	1.474968	1	115.156517	1	170.163895	1	6	1	2
H	1.118404	1	109.610477	1	-64.564886	1	16	6	7
H	1.116847	1	111.324508	1	174.220590	1	16	6	7
C	1.572621	1	110.558659	0	53.537301	0	16	6	7
O	1.260176	1	116.947977	1	25.815024	1	19	16	7
O	1.250219	1	115.394884	1	-153.956270	1	19	16	7
0	0.000000	0	0.000000	0	0.000000	0	0	0	0

METHYL MALONATE, HYDROGEN BOND ON CARBONYL FROM MENH3+
TRANSITION STATE FOR DECARBOXYLATION, -123.6 KCAL/MOL AM1

O	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.410094	1	0.000000	0	0.000000	0	1	0	0
H	1.118536	1	111.132300	1	0.000000	0	2	1	0
H	1.118706	1	104.578447	1	119.503304	1	2	1	3
H	1.120046	1	110.574853	1	-121.848193	1	2	1	3
C	1.392864	1	117.897733	1	54.716258	1	1	2	3
O	1.264274	1	114.697170	1	0.366161	1	6	1	2
H	1.785036	1	132.951096	0	25.241995	0	7	6	1
N	1.061658	1	131.415604	0	168.219380	0	8	7	6
H	1.021227	1	106.703203	1	52.536751	1	9	8	7

H	1.024641	1	105.859398	1	-62.712770	1	9	8	7
C	3.881809	1	137.870330	1	24.274125	0	7	6	1
H	1.122268	1	110.218503	1	178.771808	1	12	9	8
H	1.122998	1	109.425313	1	-60.847626	1	12	9	8
H	1.123807	1	109.252767	1	58.587850	1	12	9	8
C	1.403434	1	115.197880	1	167.966774	1	6	1	2
H	1.099835	1	115.852212	1	-45.954576	1	16	6	7
H	1.095637	1	118.153553	1	170.906837	1	16	6	7
C	2.067194	1	110.558659	0	53.537301	0	16	6	7
O	1.214618	1	108.085664	1	14.304628	1	19	16	7
O	1.210171	1	102.193874	1	-165.182363	1	19	16	7
0	0.000000	0	0.000000	0	0.000000	0	0	0	0

METHYL MALONATE (MALONYL-ACP MODEL), CARBONYL OXYGEN PROTONATED
HYDRATED WITH 5 WATER MOLECULES, -470.1 KCAL/MOL AM1

O	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.439680	1	0.000000	0	0.000000	0	1	0	0
H	1.116351	1	109.709394	1	0.000000	0	2	1	0
H	1.121380	1	102.076378	1	117.877670	1	2	1	3
H	1.117873	1	110.115572	1	-124.018809	1	2	1	3
C	1.324742	1	121.298099	1	58.348763	1	1	2	3
O	1.326054	1	121.043329	1	-1.782458	1	6	1	2
H	0.979584	1	113.415575	1	21.874745	1	7	6	1
C	1.471170	1	119.200282	1	-176.670622	1	6	1	2
H	1.120341	1	109.936803	1	-10.107634	1	9	6	7
H	1.130284	1	109.355143	1	-131.226876	1	9	6	7
C	1.562460	1	112.025527	1	109.772656	1	9	6	7
O	1.257885	1	118.875236	1	8.740269	1	12	9	7
O	1.253897	1	115.649338	1	-171.758448	1	12	9	7
H	2.321927	1	85.389418	1	163.329060	1	13	12	9
O	3.172440	1	79.502116	1	171.016133	1	13	12	9
H	3.138269	1	70.335277	1	-173.478467	1	13	12	9
H	2.292964	1	129.169391	1	-110.183149	1	14	12	9
O	2.558093	1	147.393541	1	-91.050923	1	14	12	9
H	2.226847	1	168.098613	1	-111.529217	1	14	12	9
H	2.110088	1	123.992630	1	-16.209114	1	13	12	9
O	2.427825	1	118.144426	1	10.186836	1	13	12	9
H	2.100827	1	126.291329	1	36.014330	1	13	12	9
H	3.379806	1	123.726256	1	-45.330791	1	14	12	9
O	3.025175	1	111.290790	1	-33.787069	1	14	12	9
H	2.092033	1	114.388834	1	-28.633755	1	14	12	9
H	3.500475	1	107.277720	1	54.936088	1	9	6	1
O	0.963401	1	63.711118	1	-26.331803	1	27	9	6
H	0.963990	1	103.240060	1	-100.987566	1	28	27	1
0	0.000000	0	0.000000	0	0.000000	0	0	0	0

METHYL MALONATE, HYDROGEN BOND ON CARBONYL
REACTANT FOR DECARBOXYLATION, -311.1 KCAL/MOL AM1

O	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.421559	1	0.000000	0	0.000000	0	1	0	0
H	1.117697	1	110.539605	1	0.000000	0	2	1	0
H	1.117482	1	104.167534	1	119.484822	1	2	1	3
H	1.117417	1	109.893782	1	-121.537132	1	2	1	3
C	1.375476	1	117.161449	1	58.290782	1	1	2	3
O	1.240513	1	116.111990	1	-10.124599	1	6	1	2
H	2.073821	1	131.720497	1	30.164040	1	7	6	1
O	0.972450	1	130.500763	1	-164.215614	1	8	7	6
C	2.854988	1	173.752757	1	39.061784	1	7	6	1
O	1.233557	1	113.830558	1	179.026305	1	10	9	8
H	1.107556	1	127.571843	1	179.839060	1	10	11	9
C	1.480244	1	113.947312	1	171.461793	1	6	1	2
H	1.117866	1	108.910318	1	-37.906747	1	13	6	7
H	1.117623	1	110.483626	1	-158.445982	1	13	6	7
C	1.561266	1	112.295695	1	81.874463	1	13	6	7
O	1.257110	1	118.435727	1	38.316243	1	16	13	7
O	1.258354	1	115.253112	1	-142.413816	1	16	13	7
0	0.000000	0	0.000000	0	0.000000	0	0	0	0

METHYL THIOACETATE + ENTHIOLATE
TRANSITION STATE OF CONDENSATION, -90.2 KCAL/MOL MNDO

S	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.716825	1	0.000000	0	0.000000	0	1	0	0
H	1.105971	1	113.294604	1	0.000000	0	2	1	0
H	1.108848	1	107.422313	1	119.324187	1	2	1	3
H	1.108336	1	112.008584	1	-122.469345	1	2	1	3
C	1.764345	1	112.422457	1	52.139991	1	1	2	3
O	1.236818	1	119.831924	1	18.960100	1	6	1	2
C	1.533586	1	109.572693	1	167.816575	1	6	1	2
H	1.107028	1	111.820786	1	-42.626173	1	8	6	7
H	1.106679	1	112.753353	1	-164.797065	1	8	6	7
H	1.113134	1	109.176756	1	75.979095	1	8	6	7
C	2.236064	1	96.657619	1	-90.106294	1	6	1	2
H	1.089851	1	96.125149	1	50.613778	1	12	6	1
H	1.091258	1	96.184369	1	-63.456674	1	12	6	1
C	1.426145	1	117.919408	1	142.564485	1	12	13	14
O	1.241123	1	127.645360	1	-26.807779	1	15	12	14
S	1.753748	1	113.640616	1	153.007698	1	15	12	14
C	1.716550	1	112.005588	1	164.413940	1	17	15	12
H	1.108100	1	112.662284	1	66.636882	1	18	17	15
H	1.108889	1	107.538172	1	-174.802833	1	18	17	15
H	1.107125	1	112.799885	1	-55.912916	1	18	17	15
0	0.000000	0	0.000000	0	0.000000	0	0	0	0

METHYL THIOLACETATE + ENTHIOLATE
PRODUCT OF CONDENSATION, -108.1 KCAL/MOL MNDO

S	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.703792	1	0.000000	0	0.000000	0	1	0	0
H	1.108866	1	112.929621	1	0.000000	0	2	1	0
H	1.111531	1	108.584342	1	118.835942	1	2	1	3
H	1.109086	1	112.920855	1	-122.359466	1	2	1	3
C	1.902457	1	113.282340	1	62.080626	1	1	2	3
O	1.275524	1	112.801178	1	-5.040258	1	6	1	2
C	1.566221	1	101.465435	1	118.876955	1	6	1	2
H	1.110486	1	111.110759	1	-60.936177	1	8	6	7
H	1.108647	1	112.886622	1	178.467619	1	8	6	7
H	1.108967	1	111.133209	1	57.856364	1	8	6	7
C	1.604277	1	99.513331	1	-126.649415	1	6	1	2
H	1.112995	1	108.563814	1	52.581803	1	12	6	1
H	1.110961	1	110.035338	1	-63.323948	1	12	6	1
C	1.520704	1	107.555432	1	116.643749	1	12	13	14
O	1.223523	1	125.478744	1	146.657513	1	15	12	14
S	1.736393	1	114.008906	1	-30.004624	1	15	12	14
C	1.720776	1	111.856676	1	-171.260788	1	17	15	12
H	1.106440	1	112.643043	1	55.438370	1	18	17	15
H	1.107806	1	107.092539	1	174.308343	1	18	17	15
H	1.107384	1	112.316089	1	-67.147635	1	18	17	15
0	0.000000	0	0.000000	0	0.000000	0	0	0	0

METHYL ACETATE + ENOLATE
CHARGE-DIPOLE COMPLEX, -198.0 KCAL/MOL AM1

O	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.422027	1	0.000000	0	0.000000	0	1	0	0
H	1.116703	1	110.316893	1	0.000000	0	2	1	0
H	1.117513	1	104.094850	1	119.516719	1	2	1	3
H	1.116992	1	109.974666	1	-121.182616	1	2	1	3
C	1.376915	1	116.900371	1	55.489507	1	1	2	3
O	1.235869	1	116.548050	1	6.094541	1	6	1	2
C	1.480652	1	113.370086	1	-174.574750	1	6	1	2
H	1.115323	1	109.589474	1	26.475873	1	8	6	7
H	1.134344	1	110.366747	1	-92.672485	1	8	6	7
H	1.115529	1	111.015973	1	148.502052	1	8	6	7
C	4.539838	1	98.029761	1	-114.903888	1	6	1	2
H	1.087121	1	31.292337	1	88.749890	1	12	6	1
H	1.086073	1	146.881979	1	57.521901	1	12	6	1
C	1.369350	1	119.803727	1	179.645192	1	12	13	14
O	1.269312	1	132.354835	1	-179.879331	1	15	12	14
O	1.418660	1	114.751198	1	-0.036326	1	15	12	14
C	1.408577	1	115.960787	1	-179.933639	1	17	15	12
H	1.117826	1	110.946532	1	60.473386	1	18	17	15
H	1.119548	1	104.930363	1	179.864734	1	18	17	15

```
 H    1.117982  1  110.919685  1  -60.765565  1  18 17 15
 O    0.000000  0    0.000000  0    0.000000  0   0  0  0
```

METHYL ACETATE + ENOLATE
TRANSITION STATE FOR CONDENSATION FROM CHARGE-DIPOLE COMPLEX,
-192.2 KCAL/MOL AM1

```
 O    0.000000  0    0.000000  0    0.000000  0   0  0  0
 C    1.412214  1    0.000000  0    0.000000  0   1  0  0
 H    1.119462  1  109.243708  1    0.000000  0   2  1  0
 H    1.119114  1  104.986266  1  118.546609  1   2  1  3
 H    1.117623  1  111.260267  1 -121.753292  1   2  1  3
 C    1.409027  1  115.932873  1   75.570304  1   1  2  3
 O    1.259945  1  115.024112  1   35.116383  1   6  1  2
 C    1.505793  1  109.661932  1 -178.201034  1   6  1  2
 H    1.114743  1  108.780211  1  -39.918073  1   8  6  7
 H    1.116443  1  110.434973  1 -160.102397  1   8  6  7
 H    1.116581  1  109.374400  1   79.500942  1   8  6  7
 C    2.031669  1  101.468540  1  -75.434603  1   6  1  2
 H    1.097056  1   96.519762  1 -166.355823  1  12  6  1
 H    1.096453  1   97.113095  1   76.766434  1  12  6  1
 C    1.421380  1  115.164088  1  141.454939  1  12 13 14
 O    1.249529  1  131.487816  1  169.164666  1  15 12 14
 O    1.398920  1  114.187388  1  -12.960837  1  15 12 14
 C    1.414448  1  116.524921  1 -171.491973  1  17 15 12
 H    1.117049  1  110.694839  1   56.117035  1  18 17 15
 H    1.118284  1  104.663185  1  175.721390  1  18 17 15
 H    1.117757  1  110.392128  1  -65.047373  1  18 17 15
 O    0.000000  0    0.000000  0    0.000000  0   0  0  0
```

METHYL ACETATE + ENOLATE
PRODUCT OF CONDENSATION, -199.9 KCAL/MOL AM1

```
 O    0.000000  0    0.000000  0    0.000000  0   0  0  0
 C    1.398900  1    0.000000  0    0.000000  0   1  0  0
 H    1.122409  1  109.757322  1    0.000000  0   2  1  0
 H    1.120961  1  106.240376  1  118.030638  1   2  1  3
 H    1.119249  1  111.805421  1 -121.730762  1   2  1  3
 C    1.471247  1  115.534707  1   85.419856  1   1  2  3
 O    1.297668  1  109.986668  1   36.902302  1   6  1  2
 C    1.537425  1  103.773056  1  162.766943  1   6  1  2
 H    1.113956  1  108.729393  1  -58.637782  1   8  6  7
 H    1.114031  1  110.928986  1 -179.197858  1   8  6  7
 H    1.113767  1  109.085335  1   60.312407  1   8  6  7
 C    1.592759  1  106.225088  1  -83.403185  1   6  1  2
 H    1.116792  1  106.859954  1 -169.187607  1  12  6  1
```

H	1.117722	1	106.329331	1	73.449673	1	12	6	1
C	1.479715	1	108.631431	1	120.561633	1	12	13	14
O	1.237978	1	131.124372	1	162.694994	1	15	12	14
O	1.383084	1	113.261691	1	-14.962237	1	15	12	14
C	1.419069	1	116.933165	1	-174.114868	1	17	15	12
H	1.116811	1	110.424977	1	55.476836	1	18	17	15
H	1.117578	1	104.374130	1	175.022651	1	18	17	15
H	1.117416	1	110.134581	1	-65.739041	1	18	17	15
0	0.000000	0	0.000000	0	0.000000	0	0	0	0

MALONATE ESTER + METHYL AMMONIUM ION + METHYL ACETATE
COMPLEX, -240.4 KCAL/MOL AM1

XX	0.000000	0	0.000000	0	0.000000	0	0	0	0
O	1.043215	1	0.000000	0	0.000000	0	1	0	0
C	1.425681	1	88.079845	1	0.000000	0	2	1	0
H	1.116864	1	110.292107	1	-60.853607	1	3	2	1
H	1.118876	1	110.103310	1	61.997164	1	3	2	1
H	1.119381	1	104.059443	1	-179.680706	1	3	2	1
C	1.386110	1	117.767871	1	0.000000	0	2	3	1
C	1.480841	1	120.069121	1	7.977457	1	7	2	3
O	1.234314	1	110.265362	1	-172.815430	1	7	2	3
H	1.118406	1	109.597601	1	68.477430	1	8	7	2
H	1.119336	1	109.674974	1	-171.549992	1	8	7	2
H	1.119805	1	110.583012	1	-51.925054	1	8	7	2
C	3.678248	1	94.078885	1	-77.724816	1	7	9	2
H	1.115842	1	58.234614	1	-17.810624	1	13	7	8
H	1.114597	1	82.753028	1	-137.882899	1	13	7	8
C	1.476830	1	75.303161	1	108.277063	1	13	7	8
O	1.250373	1	128.752725	1	88.751389	1	16	13	7
O	1.361615	1	115.040570	1	-90.445383	1	16	13	7
C	1.428764	1	117.156404	1	9.524380	1	18	16	17
H	1.116223	1	110.273844	1	53.128590	1	19	18	16
H	1.116288	1	109.667491	1	-68.384845	1	19	18	16
H	1.117939	1	103.447803	1	172.448188	1	19	18	16
C	1.592318	1	165.000002	0	0.000000	0	13	7	8
O	1.252156	1	114.785398	1	77.815034	1	23	13	16
O	1.248098	1	129.404298	1	-177.018699	1	23	24	13
H	2.120075	1	109.282689	1	17.991983	1	9	7	2
N	1.031582	1	139.968946	1	37.108715	1	26	9	7
H	1.045875	1	100.627324	1	105.714146	1	27	9	7
H	1.026423	1	107.967539	1	-116.035089	1	27	26	28
C	1.467055	1	111.035620	1	122.361936	1	27	28	29
H	1.121976	1	109.933648	1	-176.579478	1	30	27	28
H	1.123291	1	109.457170	1	-56.237352	1	30	27	28
H	1.123361	1	109.147937	1	63.348500	1	30	27	28
0	0.000000	0	0.000000	0	0.000000	0	0	0	0

MALONATE ESTER + METHYL AMMONIUM ION + METHYL ACETATE
TRANSITION STATE FOR DECARBOXYLATION, -235.2 KCAL/MOL AM1

XX	0.000000	0	0.000000	0	0.000000	0	0	0	0
O	1.038548	0	0.000000	0	0.000000	0	1	0	0
C	1.427431	1	87.512182	0	0.000000	0	2	1	0
H	1.116541	1	109.857610	1	-66.320246	1	3	2	1
H	1.117629	1	110.446814	1	56.397862	1	3	2	1
H	1.119212	1	103.944879	1	174.932586	1	3	2	1
C	1.377294	1	117.654724	1	0.000000	0	2	3	1
C	1.482522	1	121.089927	1	1.801184	1	7	2	3
O	1.237475	1	110.738261	1	-177.547974	1	7	2	3
H	1.117079	1	110.626630	1	54.109560	1	8	7	2
H	1.118365	1	109.675137	1	174.822001	1	8	7	2
H	1.121343	1	109.254593	1	-66.226184	1	8	7	2
C	3.678248	0	99.319101	1	-64.915710	1	7	9	2
H	1.097656	1	69.705465	1	-33.099359	1	13	7	8
H	1.095578	1	94.337583	1	-149.334783	1	13	7	8
C	1.421635	1	76.826037	1	92.976964	1	13	7	8
O	1.260173	1	129.511777	1	100.217386	1	16	13	7
O	1.377791	1	115.416393	1	-79.190509	1	16	13	7
C	1.424570	1	116.597840	1	11.944811	1	18	16	17
H	1.116346	1	110.375618	1	54.240441	1	19	18	16
H	1.116632	1	109.935013	1	-67.117874	1	19	18	16
H	1.118003	1	103.783600	1	173.642247	1	19	18	16
C	1.950000	1	165.000002	0	0.000000	0	13	7	8
O	1.219557	1	106.538661	1	93.641808	1	23	13	16
O	1.217127	1	145.307708	1	-175.449560	1	23	24	13
H	2.066979	1	115.677397	1	21.229065	1	9	7	2
N	1.028657	1	132.422490	1	27.543454	1	26	9	7
H	1.047125	1	106.576699	1	106.686879	1	27	9	7
H	1.032345	1	106.773894	1	-114.985818	1	27	26	28
C	1.464024	1	111.285494	1	121.588975	1	27	28	29
H	1.121432	1	110.411819	1	-179.346564	1	30	27	28
H	1.124347	1	109.247065	1	-58.903713	1	30	27	28
H	1.122869	1	109.333380	1	60.511434	1	30	27	28
0	0.000000	0	0.000000	0	0.000000	0	0	0	0

ENTHIOLATE + METHYL AMMONIUM ION + METHYL ACETATE
PRODUCT OF DECARBOXYLATION, REACTANT FOR CONDENSATION,
-180.1 KCAL/MOL AM1

XX	0.000000	0	0.000000	0	0.000000	0	0	0	0
O	1.026215	1	0.000000	0	0.000000	0	1	0	0
C	1.421869	1	83.238997	1	0.000000	0	2	1	0
H	1.116681	1	110.915563	1	-58.391333	1	3	2	1

H	1.118954	1	109.868860	1	64.388787	1	3	2	1
H	1.119445	1	103.949781	1	-177.221670	1	3	2	1
C	1.379564	1	117.436743	1	0.000000	0	2	3	1
C	1.485073	1	120.385831	1	10.786610	1	7	2	3
O	1.232121	1	111.177071	1	-169.871310	1	7	2	3
H	1.117552	1	110.005657	1	60.351013	1	8	7	2
H	1.118011	1	109.584576	1	-179.718882	1	8	7	2
H	1.118269	1	109.929430	1	-59.952674	1	8	7	2
C	3.678248	0	98.924654	1	-75.310817	1	7	9	2
H	1.091692	1	97.857028	1	-15.679217	1	13	7	8
H	1.092952	1	74.817573	1	-132.515393	1	13	7	8
C	1.345323	1	97.125039	1	107.312765	1	13	7	8
O	1.361105	1	128.543684	1	79.269247	1	16	13	7
O	1.373208	1	121.248334	1	-102.801817	1	16	13	7
C	1.428397	1	117.123219	1	1.718822	1	18	16	17
H	1.116018	1	110.413610	1	59.505426	1	19	18	16
H	1.116071	1	110.201874	1	-62.604303	1	19	18	16
H	1.118443	1	103.044433	1	178.635181	1	19	18	16
XX	1.592318	0	165.000002	0	0.000000	0	13	7	8
XX	1.252156	0	114.785398	0	77.815034	0	23	13	16
XX	1.248098	0	129.404298	0	-177.018699	0	23	24	13
H	2.178021	1	110.623571	1	19.883325	1	9	7	2
N	1.005068	1	137.902253	1	30.621224	1	26	9	7
H	2.099616	1	94.047745	1	91.038156	1	27	9	7
H	1.002762	1	107.881087	1	-112.916316	1	27	26	28
C	1.434069	1	114.017507	1	122.232027	1	27	28	29
H	1.125227	1	114.252176	1	179.504241	1	30	27	28
H	1.122213	1	109.297863	1	-59.610758	1	30	27	28
H	1.122456	1	108.919173	1	58.784165	1	30	27	28
0	0.000000	0	0.000000	0	0.000000	0	0	0	0

ENTHIOLATE + METHYL AMMONIUM ION + METHYL ACETATE
CONDENSATION PRODUCT WITH DOUBLE H+ TRANSFER,
-206.2 KCAL/MOL AM1

XX	0.000000	0	0.000000	0	0.000000	0	0	0	0
O	1.974824	1	0.000000	0	0.000000	0	1	0	0
C	1.414883	1	137.320426	1	180.000000	0	2	1	0
H	1.117667	1	110.688506	1	-67.617263	1	3	2	1
H	1.118091	1	110.845188	1	54.940118	1	3	2	1
H	1.119774	1	104.124929	1	173.993861	1	3	2	1
C	1.425933	1	116.318893	1	0.000000	0	2	3	1
C	1.522775	1	115.059942	1	33.077747	1	7	2	3
O	1.405588	1	101.955832	1	149.424362	1	7	2	3
H	1.116131	1	109.621317	1	51.616182	1	8	7	2
H	1.115893	1	109.200856	1	170.359932	1	8	7	2
H	1.114601	1	110.151051	1	-68.993582	1	8	7	2
C	1.540742	1	112.003978	1	-117.142285	1	7	9	2

H	1.121981	1	108.786683	1	55.578539	1	13	7	8
H	1.120771	1	109.419295	1	-62.972631	1	13	7	8
C	1.493786	1	112.633315	1	175.319521	1	13	7	8
O	1.233966	1	129.295997	1	-17.518974	1	16	13	7
O	1.367686	1	112.631862	1	163.039097	1	16	13	7
C	1.428793	1	116.820321	1	-0.255855	1	18	16	17
H	1.116456	1	109.907474	1	60.852645	1	19	18	16
H	1.116417	1	109.917358	1	-60.445235	1	19	18	16
H	1.117725	1	103.623569	1	-179.806488	1	19	18	16
XX	1.592318	0	165.000002	0	0.000000	0	13	7	8
XX	1.252156	0	114.785398	0	77.815034	0	23	13	16
XX	1.248098	0	129.404298	0	-177.018699	0	23	24	13
H	0.971277	1	108.154233	1	43.214092	1	9	7	2
N	2.806837	1	150.822977	1	-56.984304	1	26	9	7
H	1.002480	1	70.370855	1	115.097967	1	27	9	7
H	1.002878	1	52.102014	1	-147.397303	1	27	26	28
C	1.431887	1	111.078486	1	122.382538	1	27	28	29
H	1.122446	1	109.041388	1	178.773179	1	30	27	28
H	1.126075	1	114.526230	1	-60.308724	1	30	27	28
H	1.122398	1	109.032861	1	60.543886	1	30	27	28
0	0.000000	0	0.000000	0	0.000000	0	0	0	0

APPENDIX 2

ARCHIVE OF TRANSITION STATE GEOMETRIES FROM CHAPTER 5

CIS-1-NITROPROPENE CATION
RHF/HE TRANSITION STATE FOR NO2 LOSS, 266.2 KCAL/MOL AM1

C	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.472561	1	0.000000	0	0.000000	0	1	0	0
C	1.310066	1	130.892582	1	0.000000	0	2	1	0
N	1.816906	1	115.211682	1	11.844782	1	3	2	1
O	1.158931	1	115.816566	1	78.199726	1	4	3	2
O	1.168146	1	107.309388	1	-103.087237	1	4	3	2
H	1.131380	1	113.603362	1	-163.782231	1	2	3	4
H	1.094209	1	145.231752	1	-169.728677	1	3	2	4
H	1.123058	1	110.483069	1	64.775317	1	1	2	3
H	1.121958	1	111.050414	1	-175.502058	1	1	2	3
H	1.122378	1	110.366934	1	-55.665578	1	1	2	3
O	0.000000	0	0.000000	0	0.000000	0	0	0	0

CIS-1-NITROPROPENE CATION
RHF/HE TRANSITION STATE FOR THE NITRO TO NITRITE
REARRANGEMENT, 268.0 KCAL/MOL AM1

C	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.446441	1	0.000000	0	0.000000	0	1	0	0
C	1.411156	1	126.767102	1	0.000000	0	2	1	0
N	1.478678	1	123.397025	1	-15.173841	1	3	2	1
O	1.151599	1	153.169802	1	-87.847290	1	4	3	2
O	1.265076	1	76.137635	1	94.360088	1	4	3	2
H	1.113172	1	115.064054	1	165.623437	1	2	3	4
H	1.118437	1	123.083732	1	-169.411905	1	3	2	4
H	1.130221	1	109.803527	1	119.694335	1	1	2	3
H	1.131032	1	109.266842	1	-125.002679	1	1	2	3
H	1.117745	1	114.834557	1	-3.245965	1	1	2	3
O	0.000000	0	0.000000	0	0.000000	0	0	0	0

CIS-1-NITROPROPENE CATION
RHF/HE TRANSITION STATE FOR H TRANSFER FORM C3 TO O,
259.0 KCAL/MOL AM1

C	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.447477	1	0.000000	0	0.000000	0	1	0	0
C	1.349242	1	123.593025	1	0.000000	0	2	1	0
N	1.482576	1	119.901972	1	7.946509	1	3	2	1
O	1.180147	1	121.468956	1	-172.173900	1	4	3	2
O	1.248468	1	118.479657	1	8.574945	1	4	3	2
H	1.117041	1	118.908619	1	-171.004935	1	2	3	4
H	1.109687	1	127.434873	1	-179.592901	1	3	2	4
H	1.124623	1	114.874983	1	78.477856	1	1	2	3
H	1.119693	1	115.698610	1	-146.509560	1	1	2	3
H	1.617934	1	114.492292	1	-1.953327	1	6	4	3
O	0.000000	0	0.000000	0	0.000000	0	0	0	0

PRODUCT OF H TRANSFER FROM C3 OF C-1-NP+ TO O
3X3 CI TRANSITION STATE FOR LOSS OF OH, 255.3 KCAL/MOL AM1

C	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.373191	1	0.000000	0	0.000000	0	1	0	0
C	1.400414	1	128.829508	1	0.000000	0	2	1	0
N	1.361015	1	127.612797	1	5.402652	1	3	2	1
O	1.137269	1	139.111242	1	-178.165153	1	4	3	2
O	2.019259	1	116.489436	1	5.759621	1	4	3	2
H	1.114726	1	112.795416	1	-175.352421	1	2	3	4
H	1.120051	1	118.841171	1	179.556314	1	3	2	4
H	1.102970	1	124.627890	1	2.756380	1	1	2	3
H	1.106078	1	119.822026	1	-178.239945	1	1	2	3
H	0.954933	1	109.620139	1	179.395724	1	6	4	3
O	0.000000	0	0.000000	0	0.000000	0	0	0	0

CIS-1-NITROPROPENE CATION
RHF/HE TRANSITION STATE FOR H TRANSFER FROM C1 TO O,
298.6 KCAL/MOL AM1

C	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.456441	1	0.000000	0	0.000000	0	1	0	0
C	1.356647	1	124.908753	1	0.000000	0	2	1	0
N	1.451867	1	140.806359	1	-40.275652	1	3	2	1
O	1.250787	1	103.501132	1	177.694610	1	4	3	2
O	1.172156	1	131.397423	1	-3.774432	1	4	3	2
H	1.119283	1	116.889250	1	147.516712	1	2	3	4
H	1.434974	1	85.698827	1	-2.154217	1	5	4	3
H	1.132306	1	107.751755	1	124.966864	1	1	2	3
H	1.127726	1	110.622908	1	-119.132544	1	1	2	3
H	1.119358	1	114.209343	1	4.240303	1	1	2	3
O	0.000000	0	0.000000	0	0.000000	0	0	0	0

CIS-1-NITROPROPENE CATION
RHF/HE TRANSITION STATE FOR LOSS OF H FROM C3,
286.5 KCAL/MOL AM1

C	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.499151	1	0.000000	0	0.000000	0	1	0	0
C	1.353726	1	111.995843	1	0.000000	0	2	1	0
N	1.480957	1	108.832258	1	-0.009507	1	3	2	1
O	1.166514	1	129.384838	1	179.987263	1	4	3	2
O	1.633695	1	99.008836	1	0.025864	1	1	2	3
H	1.101695	1	125.762584	1	180.012782	1	2	3	4
H	1.101610	1	133.944226	1	-179.981520	1	3	2	4
H	1.129910	1	117.671652	1	101.318940	1	1	2	3
H	1.537874	1	97.108168	1	179.917626	1	1	2	3
H	1.129771	1	117.649361	1	-101.301370	1	1	2	3
O	0.000000	0	0.000000	0	0.000000	0	0	0	0

END

12-86

DTIC